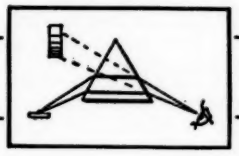
Journal of the OPTICAL SOCIETY *of* AMERICA

VOLUME 33
No. 11



NOVEMBER
1943

Visual Processes and Color Photography . . . . . R. M. EVANS 579

Herschel Effect in Relation to a Critical Size of Development Centers . . . . . G. KORNFELD AND T. H. JAMES 615

Effect of Hue on Dark Adaptation . . . . . EARL M. LOWRY 619

Mathematical Proof of Stokes' Reversibility Principle . . . . . ARCHIE I. MAHAN 621

Color Designations for Lights . . . . . KENNETH L. KELLY 627

Graphical Representation of Small Color Differences . . . . . W. D. WRIGHT 632

Letters to the Editor:

Note on the Correlation Method in Quantitative Spectrum Analysis . . . . . SAUL LEVY 637
Optical Physics in the U.S.S.R. . . . . S. FRISCH 637

Book Review 639

Announcements:

New Institute for Fundamental Research in Glass 614
New House Organ of Interest to Spectrographers 620
New Headquarters Building of the American Institute of Physics 626

Published monthly
at Prince and Lemon Sts., Lancaster, Pa.
by the
AMERICAN INSTITUTE OF PHYSICS
Incorporated
175 Fifth Avenue, New York 10, N. Y.
FOR AND UNDER THE EDITORIAL SUPERVISION OF THE
OPTICAL SOCIETY OF AMERICA

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

THE OPTICAL SOCIETY OF AMERICA

Officers of the Society

A. H. PFUND, *President*
Johns Hopkins University
Baltimore, Maryland

GEORGE R. HARRISON, *Vice President*
Massachusetts Institute of Technology
Cambridge 39, Massachusetts

ARTHUR C. HARDY, *Secretary*
Massachusetts Institute of Technology
Cambridge 39, Massachusetts

W. W. GRAEPER, *Treasurer*
Bausch & Lomb Optical Company
Rochester 2, New York

The Board of Directors consists of the above officers; the Editor of the Journal; the Past President; and five directors-at-large.

A. G. WORTHING, *Past President*
University of Pittsburgh
Pittsburgh, Pennsylvania

GEORGE R. HARRISON, *Editor*
Massachusetts Institute of Technology
Cambridge 39, Massachusetts

Directors-at-Large

WALLACE R. BRODE
Ohio State University
Columbus, Ohio

SELIG HECHT
Columbia University
New York 27, New York

D. B. JUDD
National Bureau of Standards
Washington, D. C.

D. L. MACADAM
Eastman Kodak Company
Rochester 4, New York

R. A. SAWYER
Naval Proving Ground
Dahlgren, Virginia

Rochester, New York, Section

F. W. PAUL, *President*
P. W. STEVENS, *Vice President*
A. F. TURNER, *Secretary*
F. H. PERRIN, *Treasurer*

Councillors:

F. C. FAIRBANKS
J. H. MCLEOD
W. T. PERKINS
D. L. WILLIAMS
J. H. WEBB (*ex officio*)

Patrons

AMERICAN OPTICAL COMPANY BAUSCH & LOMB OPTICAL COMPANY
EASTMAN KODAK COMPANY

Corporation Members

AMERICAN CYSTOSCOPE MAKERS, INC., New York, New York
AMERICAN OPTICAL COMPANY, Southbridge, Massachusetts
BAUSCH & LOMB OPTICAL COMPANY, Rochester, New York
BELL & HOWELL COMPANY, Chicago, Illinois
BELL TELEPHONE LABORATORIES, INC., New York, New York
BENDIX AVIATION CORPORATION, New York, New York
CENTURY LIGHTING, INC., New York, New York
CORNING GLASS WORKS, Corning, New York
E. I. DUPONT DE NEMOURS & COMPANY, INC., Chemical Department, Experimental Station, Wilmington, Delaware
EASTMAN KODAK COMPANY, Rochester, New York
FAIRCHILD AVIATION CORPORATION, Jamaica, New York
FISH-SCHURMAN CORPORATION, New York, New York
GAERTNER SCIENTIFIC CORPORATION, Chicago, Illinois
W. & L. E. GURLEY, Troy, New York
INTERCHEMICAL CORPORATION, New York, New York
KEUFFEL & ESSER COMPANY, Hoboken, New Jersey
NATIONAL CARBON COMPANY, Cleveland, Ohio
PERKIN-ELMER CORPORATION, Glenbrook, Connecticut
POLAROID CORPORATION, Cambridge, Massachusetts
PORTMAN MACHINE TOOL COMPANY, New Rochelle, New York
SIMPSON OPTICAL MANUFACTURING COMPANY, Chicago, Illinois
SPENCER LENS COMPANY, Buffalo, New York
ARTHUR H. THOMAS COMPANY, Philadelphia, Pennsylvania
ZENITH OPTICAL COMPANY, Huntington, West Virginia

THE OPTICAL SOCIETY OF AMERICA was organized in 1916 "*to increase and diffuse the knowledge of optics in all its branches, pure and applied, to promote the mutual interests of investigators of optical problems, of designers, manufacturers and users of optical instruments and apparatus of all kinds and to encourage cooperation among them.*" It aims to serve the interests of all who are engaged in any branch of optics from fundamental research to the manufacture of optical goods.

MEMBERSHIP

Regular Members: "Any person who has contributed materially to the advancement of optics is eligible for regular membership."

Associate Members: "Any person who is interested in optics is eligible for associate membership."

Corporate Members: "Any corporation desirous of encouraging the work of the Society is eligible for corporation membership."

Patrons: The payment, within any one year, of $1000 or more toward the expenses of the Society by any corporation or individual entitles such corporation or individual to the rank of patron.

The Society cordially invites to membership all who are interested in any branch of optics, either in research, in instruction, in optical or illuminating engineering, in the manufacture and distribution of optical goods of all kinds or in physiological and medical optics.

ANNUAL DUES

Regular members, $8.50; Associate members, $6.00; Corporate members, $50.00; Patrons, $1000.00. These dues include subscription to the Society's official journal, the *Journal of the Optical Society of America.*

CHANGES OF ADDRESS

All changes of address of members of the Optical Society of America should be communicated to the Secretary, Professor Arthur C. Hardy, Massachusetts Institute of Technology, Cambridge 39, Massachusetts.

Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 33, NUMBER 11 NOVEMBER, 1943

Visual Processes and Color Photography*

R. M. EVANS
Eastman Kodak Company, Rochester, New York

I. BRIGHTNESS RELATIONS

Introduction

IN the theory of black-and-white photography, very careful consideration has been given to the relationship required between the luminance** of areas in the object and that of the corresponding areas in the reproduction. Great advances in photography have been made by a combination of theoretical and empirical methods. An exact theoretical correspondence of all luminances in object and print has been found not only impossible but in some cases undesirable. An attempt to reproduce the appearance of a three-dimensional colored object involves more than simply abstracting the luminance relations and excluding the chromaticity.

In many cases, chromaticity is as important a factor in the contrast of a scene as are the actual luminance differences. In the extreme case, two surfaces may be identical in luminance but have sharply contrasting colors. This problem is solved by the use of "contrast filters" which enhance or reduce the effective luminance of the different areas of a particular scene by modifying the relation between energy distribution and reproduction density. Experience has shown that these considerations are of great importance only in those cases where desired detail would otherwise be lost, as in the photography of furniture and the like. The theory of tone reproduction can be developed for any given relation between chromaticity in the object and density in the reproduction, and since these relations all become identical for the reproduction of a subject consisting of white, gray, and black areas only, this is the aspect which is usually discussed.

The aim of black-and-white photography is to produce either a pictorial or an instructional reproduction of an object in a scene. The absence of chromaticity precludes the possibility of obtaining a perfect reproduction in the sense that it might be mistaken for the object itself. Because there is no expectation that the result will look exactly like the object, it is not necessary to consider very carefully all aspects of the differences. In the tone-reproduction theory as developed by L. A. Jones and others,[1] little emphasis has very properly been placed therefore, on the subjective phases of the visual processes involved and attention has been concentrated on the prob-

* Paper presented at the meeting of the Optical Society of America held in New York City, March 5–6, 1943. Communication No. 944 from the Kodak Research Laboratories.

** Throughout this paper, "luminance" will be used to indicate the psychophysical intensity aspect of radiant energy as defined by the O. S. A. colorimetry committee (J. Opt. Soc. Am. **27**, 207 (1937)). "Brightness" will be used to refer to the subjective appreciation of luminance, and "lightness" will refer to the subjective appreciation of the luminance of reflecting surfaces.

Readers of *The Theory of Photography* by C. E. K. Mees should bear in mind that in Chapter XX of that work the recommendations then current were followed. In the former notation *brilliance* was used for the present *brightness* and *brightness* was used for the present *luminance*.

[1] C. E. K. Mees, *The Theory of Photography* (Macmillan, New York, 1942), Chapter XX.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

lem of luminance rather than brightness relationships. This work has shown that of all possible relations between the luminance of the object and that of the corresponding area in the print, a linear relationship gives the most acceptable result. None of the conclusions which follow affects the validity of this fact.

Color photography is theoretically capable of reproducing all the visible characteristics of a given scene. Except for physical restrictions on the range of luminance and chromaticity which can be reproduced simultaneously in one picture, it is not impossible to obtain a reproduction which may be mistaken for the object itself. This possibility has led many workers to assume that this is the goal of color photography and that if the reproduction were to meet the requirement that every point of the final picture shall match the corresponding point of the object in chromaticity and luminance, the picture would look exactly like the object. Experience indicates that the latter view is merely an assumption since in a particular case this condition is neither necessary nor sufficient to produce this result. It may be the only *possible* assumption on which to base a theory for the *general* case.

It is the purpose of this article to consider the subjective phases of color photography in their bearing on this particular point, in the hope that it will provide a more secure footing for theory as well as practice. Present knowledge of the subjects involved will be reviewed and their relative importance indicated by the consideration of the effects observed in the actual operation of color processes. Part I considers only brightness relations. Contrast and adaptation phenomena will be discussed separately.

Luminance *vs.* Brightness in Photography

One of the basic assumptions which underlies all theories of photographic color processes is that a photograph *can* be made which *looks like* the subject. It is usually postulated also that if each area of the reproduction were to match each area of the subject, such a photograph would be obtained. In cases where no such postulate is made it is at least assumed that each area in the reproduction must bear some definite relation to that of the subject.

Restricted to luminance, these assumptions imply that there is a definite expressible relation between the luminance which evokes a given brightness at any point in the scene and the luminance required to evoke the same brightness in the corresponding point of the reproduction.

For the reproduction to appear as a positive rather than a negative, at least the greater part of the brightnesses must come in the right order in the reproduction. Since the assumption states that there is an expressible relation, it would follow that *the order* of *all* brightnesses would be correct. Hence, if in the subject one brightness is greater, equal to, or less than another it will so appear in the print. Anticipating the results to be discussed, it was found that such a general relation does not exist except in special cases. Unless a scene is suitably illuminated when the picture is taken, any object in it which appears lighter than another may reproduce as equally light or even considerably darker. In other words, the *order* of brightnesses in the reproduction does *not* necessarily appear the same as it does in the subject.

At first sight this may seem difficult to believe. It is true because any satisfactory photographic process produces a fairly accurate record of the relative luminances of a subject but there is no direct relation in vision between relative luminance and relative brightness. One is tempted to say there is no relation at all. The brightness evoked by a given luminance depends on the circumstances at the time of viewing. Since these circumstances are, in general, different according to whether the subject or the print is viewed, the effect is different in the two cases and the order of the two brightnesses may actually be inverted.

In order to understand the failure of photography faithfully to reproduce the brightness of the original, it is necessary to review the lack of correspondence between luminance and brightness in vision. The fundamental question here, however, is why the print does not have the same brightnesses as the subject even when the luminances are exactly matched.

Brightness Constancy in Nature

Various "constancy phenomena" are well established in the psychology of visual processes. They may be summarized conveniently in the

statement that objects in nature tend to be perceived by the mind more nearly as they exist in space than as the two-dimensional projection formed by the eye.

"Size constancy" is illustrated by the fact that people at a distance do not *look* smaller than those close at hand. The process by which this occurs is one of comparison, combined with appreciation of perspective. The magnitude of these objective mental corrections is large. A person at 100 yards forms a retinal image which is one-tenth the size formed at 10 yards and yet he looks scarcely any smaller—he simply looks farther away.

"Shape constancy" is frequently described by pointing out that a penny or a hoop seen at any angle except directly edge-on is perceived as a circle rather than an ellipse. Really this is a single example of the general interpretation of space as three-dimensional through familiarity with the effect of perspective. For relatively near objects, the interpretation is believed to take place through cortical fusion of differing retinal images in the two eyes. For more distant objects, it must occur through perspective. Occasionally the perception is noticeably wrong, as in the apparent meeting of railroad tracks in the distance, but normally it is sufficiently accurate to be of great aid in daily life.

Other phenomena of vision may be considered as constancy effects. Thus, we might speak of a "position constancy" of objects in a room relative to a moving observer or of a "motion constancy" of objects which are seen to be moving at the same speed although they are at various distances from the observer.

These phenomena, which are apparent even to the casual observer, are probably all learned reactions and may be understood readily by their utility to the individual and their relation to known laws of perspective and of relative motion. Perspective relations can be well reproduced by photography and it is not surprising therefore to find that, by the proper choice of the point of view and the focal length of the lens, the size and shape of objects may be reproduced satisfactorily in two dimensions. To secure the effect of depth for near objects requires stereoscopic reproduction. Relative motion phenomena in which the observer is supposed to be moving are not usually well represented by motion pictures because the observer does not have the sensations accompanying motion.

The extension of the constancy principle to brightness and to chromaticity which has been made in recent years[2] rests on somewhat less secure ground. A very close parallel can be drawn between the constancy of the chromaticity and the brightness of an object and the constancy of its size and shape. For practical purposes, this is undoubtedly the most satisfactory explanation of the discrepancy between the perceived effect and the actual physical phenomenon.

On the other hand, it has not been demonstrated that it is necessary to assume psychological processes in order to explain the observed facts. In photography we are concerned with the reproduction of the visual effects since they are important factors in the recognition of objects. It does not matter greatly which approach is used in explaining them so long as it will satisfactorily include all the facts.

If a white or a light gray surface is placed in a shadow and a similar surface is placed in full illumination, the observer will have no difficulty in recognizing their identity. White surfaces will continue to look white under both conditions and the lightness of the gray will not appear to be appreciably different. Many manifestations of this fundamental fact may be demonstrated and the magnitude of the luminance difference between two surfaces, both of which may be perceived as white, is startling to the uninitiated. It is not difficult to produce a luminance ratio of 100 to 1 and yet have both surfaces readily recognizable as white. To say that their brightness is the same under these conditions would be an exaggeration since the shadowed white is always noticeably darker. It is, however, *much* brighter than a gray which reflects the same amount of light but is placed in the illuminated part of the field. It should be noted that brightness and lightness must be distinguished in such a case since the object which appears less bright is also perceived as lighter than the other. An

[2] D. Katz, *The World of Color* (Kegan Paul, Trench, Trubner, and Co., London, 1935); K. Koffka, *Principles of Gestalt Psychology* (Harcourt, Brace and Co., N. Y., 1935); A. Gelb in Bethe's *Handbuch der normalen und pathologischen Physiologie* (Julius Springer, Berlin, 1929); E. R. Jaensch, Zeits. f. Sinnesphysiol. **52**, 165 (1921).

extension of this demonstration is even more startling. Under the above conditions a good black in the illuminated area may actually reflect more light to the eye (i.e., have a higher *luminance*) than the white surface in the shadow. In this case the black has very low brightness and lightness compared to the white. There is, therefore, no direct relation between brightness and luminance in a non-uniformly lighted scene. In uniform illumination, which is rare and should be considered as a special case, lightness, brightness, and luminance always occur in the same order.

An illuminated scene has two characteristics by which it is seen. A certain percentage of the light which falls on the objects in the scene is reflected and a certain quantity of light is reflected from the scene as a whole. The first of these distinguishes one object from the objects in its immediate vicinity, and the second is the actual stimulus for the eye of the observer. If these characteristics are called reflectance and luminance, respectively, then it may be stated that the perceived lightness of a gray surface having a given reflectance at a given luminance in a non-uniformly lighted scene will always fall between that given by the same reflectance at the high illumination and the same luminance at the low illumination. Thus a medium gray surface in shadow will appear to match a gray in the light which has a lower reflectance but a higher luminance. This is the equivalent of saying that a shadowed (or less strongly illuminated) object tends to be judged by its reflectance rather than by its luminance. This is not strange when it is realized that it can only be distinguished from its background *at all* if it has a different reflectance.

The degree to which the brightness depends on the one factor or the other is determined almost entirely by circumstances at the time of viewing. The most important of these is the extent to which it is apparent to the observer that the illumination is different on the two surfaces being compared. Scarcely less important, and in many cases essential is a continuity of background so that the surfaces are either seen against a surface which is common to both or against surfaces which do have this relation to a common background.

The role of the observer appears not to have been thoroughly investigated. If conditions are such that the observer can make a quantitative setting, different observers will make matches varying all the way from equations of reflectance to equations of luminance. Some observers can, on request, make two settings, one of which approaches a reflectance match and the other a luminance match. Children and untrained observers uniformly tend to make reflectance matches.[3]

It is this latter fact, namely, that *lack* of training and experience tends to produce the closest approach to brightness constancy that sets this phenomenon apart from that of size and shape constancy in which the reverse is true. A psychological effect might be expected to be more complete in an adult.

A somewhat better picture of the subject and one which is directly applicable to photography can be obtained by a more careful consideration of what is meant by the terms white, gray, and black. The variables on which they depend give at least a partial clue to the results of lightness matches.

There has been much discussion in the literature concerning the nature of gray. It has generally been assumed that one or another type of retinal or cortical mechanism is involved. For the present purpose it is not necessary to take any stand on this subject but it is essential to consider carefully the facts insofar as they are known.

It has been noted above that brightness is not directly related to luminance. A moment's consideration will show that for the same reason there is no relation between luminance and the characteristics white, gray, and black. This follows from the fact that a black surface in the light can be more luminous but darker than a white surface seen at the same time in shadow. Furthermore, this is a general phenomenon of which examples may be multiplied indefinitely. Two neutral surfaces of any reflectance can be made to have any desired lightness with respect to each other if each of them and their surroundings are separately and properly lighted. An intensely illuminated "black" appears white if

[3] D. Katz, reference 2; R. H. Thouless, Bull. J. Psychol. **21**, 348 (1931).

there is no comparison surface, and a white may be made black or any shade of gray by simply surrounding it with a more highly illuminated white.[4]

White, gray, and black, therefore, have no physical existence in nature. They are descriptive terms arising from the perception of lightness. The physical property common to them all is nonselective absorption of the illumination. The proper term for the appearance of a particular nonselective body depends on the illumination and surroundings at the time it is seen. In particular, since a body of very low reflectance (black) will be seen as a gray or a white if viewed against a darker background, and a white as a gray or black if seen against a lighter one, *white*, *gray*, and *black* are the terms which describe the perception of relative brightnesses. In the special case of a uniformly illuminated field, they are identical with the perception of relative luminances.

In daylight and in familiar surroundings, a nonselective object with a medium reflectance is seen as nearly the same shade of gray under all conditions. The phenomenon of brightness constancy reduces to the statement that the gray produced depends on the luminance of the gray relative to its *immediate* surroundings. Surroundings which lie on the other side of a shadow edge or at a distance in non-uniform light are excluded so far as the immediate perception of white, gray, or black is concerned. To what extent a shift of the adaptation level of the eye produced by crossing the boundary or by centering the attention on one side or the other affects this phenomenon is not known. It can, of course, be stated that insofar as the eye tends to divide up the scene into regions of *uniform* illumination, perception is controlled by the relative luminances within this region. This is not the whole story, however, because in such a case it is not necessary that within that particular, small, uniformly lighted area, either a white or a black be perceived in order to recognize a gray.

The importance of these considerations for black-and-white photography does not require emphasis. Color photography is equally concerned with the reproduction of neutral areas. It is also concerned with brightness relations among the colors themselves. Little work which is directly applicable to the brightness constancy of chromatic surfaces in non-uniform illumination appears to have been published, although Helson and Judd[5] have contributed definite information in this field. A little thought or a few simple experiments will quickly convince the curious that the phenomenon is not at all limited in its scope to even approximately nonselective reflectors. It is, in fact, not even restricted to reflecting bodies, being produced almost equally well by luminous areas, provided their luminance is not much in excess of that of their surroundings. Because the general case is so intimately associated with visual adaptation both for luminance and for chromaticity, further attention will not be given the subject here.

Photography of Brightness Constancy Phenomena

The photography of brightness constancy phenomena is usually considered impossible by the psychologists working in this field. Katz[2] devotes a section of his book to the subject. He bases his conclusion that "photographs can never give perfectly natural impressions" on the three factors of limited visual angle, decreased size of the reproduction, and monocular viewpoint. He states that "photography provides only a unidimensional series of achromatic colors" and explains this on the basis that photography cannot (at least to the same extent) "utilize certain determinants which have measurable influence upon the character of normal perception." MacLeod,[6] in his exceptionally thorough and careful study of the subject, used photographs of his apparatus in at least one case to determine whether or not a true luminance match actually existed. He states that "in the reproduction . . . the two sides *appeared* . . . almost exactly the same" (italics are mine). In this case it is assumed that the photograph itself shows *no* tendency towards brilliance constancy. *Readings* of the resulting densities in the repro-

[4] R. S. Woodworth, *Experimental Psychology* (Holt & Co., N. Y., 1938), p. 14.

[5] See, for example: H. Helson, J. Exp. Psychol. **23**, 439 (1938); H. Helson and V. B. Jeffers, J. Exp. Psychol. **26**, 1 (1940); D. B. Judd, J. Research Nat. Bur. Stand. **24**, 293 (March, 1940).

[6] R. B. MacLeod, Arch. Psychol. **21**, 1–101 (1932).

duction would, of course, have been a legitimate check for a luminance match. The assumption that if two densities in a photograph look the same they are the same, however, requires demonstration for the general case.

This assumption is important in photography because it underlies the theory of tone reproduction, i.e., that there should be an exact point-for-point match of the relative luminances in the subject and the reproduction. If two equal densities in a picture always appear of equal brightness, it is seen that this leads to a false reproduction, since the same statement is not true of objects in nature. The requirement is that the lightnesses evoked by the print be the same as those evoked by the subject, since only in this manner can a shaded white seem white and not gray or even black. The actual densities in the print are of secondary importance except for purposes of photographic photometry.

The identical problem exists in art in those cases in which an attempt is made to reproduce the appearance of a scene. The writer cannot agree with those psychologists who maintain that an artist must learn to overcome brightness constancy effects so that he can make the relative luminances of the objects in his painting match those in the scene. If the artist were to reproduce the exact relations of the luminances in the scene, then, in order for his painting to appear satisfactory, it should produce the same degree of brightness constancy in the observer's perceptions as the scene. What the artist can and must do is to adjust the relative luminances in his painting so that the perceived lightnesses correspond to his perception of the scene as a whole. He must, therefore, train himself to see and appreciate the naïvely observed lightnesses rather than luminances. This is true, incidentally, whether the perceived lightnesses are due to brightness constancy or to simultaneous contrast. By a peculiarly circular process of reasoning, probably due largely to Chevreul, it is frequently stated that it is necessary to be able to eliminate simultaneous contrast effects visually in order to paint a scene realistically. Aside from the practical impossibility of doing so without visual aids, an exact luminance match of two areas will not reproduce the simultaneous contrast, if, for example, the areas of the reproduction are different from those of the scene. The requirement again is that the relative luminances be so adjusted that the relative lightnesses *do* match those being reproduced.

The artist is obviously in a better position than the photographer to make arbitrary luminance corrections from point to point of a picture. The photographer, however, is not entirely helpless in this respect and a photographic reproduction does not always lack brightness constancy effects.

In order to learn the nature of brightness failures in photography and, if possible, to lay the groundwork for techniques which would give more faithful reproduction from this point of view, a brief study has been made of the problem of photographing scenes which show large differences between apparent and actual luminances.

If two pictures which are identical in every respect, except that one is darker throughout the entire scale than the other, are hung against a common background, no amount of light shining on the darker one will make it look lighter than the other. A photograph of the two pictures under these conditions, however, will frequently show the opposite effect, the darker more strongly illuminated picture reproducing as the lighter of the two. If, instead of illuminating the dark picture, a shadow is allowed to fall across the scene so that the lighter picture is in the shadow, a photograph can be taken in which the light picture still looks lighter (higher brightness) in the reproduction even though it is actually darker (lower luminance) just as in the original. (The experiment is a striking one in color photography but does not reproduce well in black and white for a reason to be discussed presently.) An underlying principle is apparent here which seems to hold for all cases in which brightness reproduction is satisfactory. The shadow obviously divides the scene into two areas of different illumination. The extent to which the illumination difference is apparent appears to control the perceptions produced. This illumination difference is shown by the continuous background.

A large wall of yellow tile across which the sun threw shadows at a convenient time of day was chosen for further experiments. Large squares of gray, white, and black papers were arranged in two series. One series was placed against the wall in shadow and the other in sunlight. Photographs

FIG. 1.

of this scene with a person included for reference were taken by single-lens and stereoscopic cameras in both color and black and white. The scene itself is illustrated in Fig. 1. The luminance of the shadow was so adjusted at the time this picture was taken that the second paper from the black end in sunlight matched the white in the shadow. The ratio of sunlight to shadow illumination was about 20 to 1. Visually the lightness of the white in shadow was very nearly the same as the white in sunlight.

Visual examination of the reproduction photographed by the various techniques led to the following conclusions:

(1) At least some brightness constancy was shown by all the photographs however they were viewed.

(2) The best reproduction of the visual effect was given by stereoscopic color photography. Distinctly the worst reproduction was given by a black-and-white print from a single-lens camera negative. Stereoscopic black-and-white reproduction was nearly but not quite as good as a single-lens color transparency viewed over an illuminator in a dark room.

(3) Various observers showed more variation in their description of the reproductions than they did of the scene itself.

(4) An attempt to learn to what extent the results were due to simple simultaneous contrast appeared to indicate that a large part of the effect in the poorest reproduction could be ascribed to this cause. The color transparency, however, had a large effect apparently caused by perceptual phenomena only.

This last conclusion is derived from tests made in the following manner. The strips carrying the series of grays were carefully cut out of the reproductions and remounted on neutral backgrounds which had the same density as the backgrounds in the picture. These backgrounds were so arranged that they covered the same areas and were arranged in the same manner as the shadows and sunlight in the original reproduction. The results obtained by a number of observers indicated that the lightnesses of the gray papers in the shadow in the original photograph were definitely higher than on the corresponding gray background. For example, if in the original photograph a given step in the sunlight appeared to match the white in the shadow, then a luminance nearly four times as great was required for the corresponding match on the gray background. This is not considered as complete proof that simultaneous contrast is not largely responsible for the effect, however. Very little is known about the quantitative laws followed by the brightness of adjacent areas. Until such laws are formulated it seems better to note simply that a subject which gave a brightness match with a luminance ratio of about 20 to 1 was reproduced so that a ratio of 16 to 1 appeared to match in the color transparency and that a simple simultaneous contrast imitation of the scene gave a match at a ratio of about 4 to 1.

Other experiments were made on other types of subjects with the same general results. In every case color was better than black and white and stereoscopic photography was better than single lens. In no case was there complete absence of the effect. The greatest effect observed (in a color transparency) had an area which was judged by all observers to be definitely lighter than another area but which actually had a density 1.5 greater (about 1/30 the luminance).

It is felt that such large differences justify the conclusion that brightness constancy plays a large part in photographs on non-uniformly lighted scenes. It also indicates the need for careful quantitative work to attempt to learn the conditions which give the best results and the

average which may be expected under given conditions.

Some of the consequences for photography are immediately apparent. The fact that the various techniques arrange themselves in the order found is the equivalent of the statement that the more *realistic* the reproduction is made the greater the effect becomes.

Tests made indicate that the requirement of realism applies also to the contrast of the reproduction. In black and white, at least, prints made at too high and too low contrast levels appeared to show less effect than those at the normal level.

The surroundings under which the reproduction was viewed played a large part, as might be expected. Color transparencies viewed in a dark room were more effective than in daylight surroundings.

Scenes with very contrasty lighting gave poor results in general. It is perhaps this, as much as a misunderstanding of the whole problem that has given rise in the literature of color photography to the ridiculous notion that the latitude of color film is so short it will only handle luminance ratios of 4 to 1. The actual fact is somewhat as follows: In an actual scene the luminance ratios of objects in uniform illumination may be as high as 50 to 1. If a shadow falls across part of this scene and the illumination in this shadow is one-fourth that of the brighter part, the total luminance ratio of the scene will be 200 to 1. This is about the scale of most color films. A shadow with one-fourth the intensity of the main light is also about the darkest which can be tolerated photographically because of failure of the brightness constancy effect in the print. The proper statement, therefore, is that the ratio of maximum to minimum *illumination* in any given scene should not be greater than 4.

Again, there is much confusion about the colors in which backgrounds reproduce. If a background is to reproduce as the same color which it appears to the photographer, it must be separately lighted to very nearly the same level as foreground objects. The brightness constancy which makes it look normal under all conditions with respect to the walls and the adjoining areas of the studio fails completely when it is seen as an unrelated area in the final picture. A gradual falling away of light from the front to the back of the scene is hardly noticeable in the studio but shows up very strongly in the reproduction.

In black-and-white work, even though the shadows are reproduced at the same relative luminance as in the scene, they appear much darker in the picture. The fault is not in the tone reproduction theory and it cannot be corrected by changing the contrast of the print. The difficulty lies in the lack of realism of black and white as a reproduction medium and the consequent loss of most of the constancy of brightness in the shadows.

The photographer can do much to improve the final result by modification of his lighting. The requirement of increased lighting for backgrounds has already been mentioned. The same principle also holds true for shadows. The principal light source must give fairly strong shadows to be convincing, that is, it must throw shadows with fairly well defined edges. Then in order that objects in this shadow may appear nearly as bright in the print as in the subject, it is necessary that the *contrast* of these shadows with respect to other illuminated areas be very greatly decreased by auxiliary illumination. This is the basis of the motion picture cameraman's rule that a set should be considered as a tank to be filled with light just as a tank is filled with water and then the principal light should be added. In the extreme case there is no principal light and no brightness constancy phenomenon. It is this which is meant when "flat lighting" is recommended for color photography.

II. COLOR ADAPTATION AND COLOR CONSTANCY

Introduction

Sensitivity adjustments of the eye play a large part in vision. For purposes of descriptive discussion these adjustments may be considered as consisting of two types, those for brightness and those for chromaticity. There is good evidence that this division also corresponds to an actual division of functions in the eye. The present article is intended as a brief review of the properties of the eye insofar as they depend on the energy distribution in the spectrum of the illuminant and the light from the objects in the field of view. It is therefore concerned with the adjustments of the eye for color. These are thought of

as occurring independently of any simultaneous adjustment for brightness. The justification for the point of view taken comes largely from the recent work of Wright[7] and of Schouten.[8] Some parts of the discussion must be considered as speculative extensions of their data but an attempt will be made to point out these parts as they occur. It is felt that this new approach to the subject may, in itself, have some usefulness in the field of physiological optics aside from the application of its conclusions to the theory of color photography.

One of the basic problems of color photography lies in the requirement that the reproduction must look sufficiently like the original subject under a wide variety of illuminants and illumination conditions. The fact that such a requirement can be imposed and that to a first approximation it has been met by successful processes implies a visual phenomenon of a remarkable nature. This phenomenon, known as "color constancy," has been variously described as purely psychological and as purely physiological. The present approach is intended to illustrate the problems which it raises for color photography. For this purpose a combination of these two approaches based on the work of Wright and of Schouten gives a description which leads to concepts useful in predicting and explaining photographic effects. No attempt will be made to make the concepts quantitative beyond indicating the nature of the variables. It is unlikely, however, that a complete theory of color photography can be successful without such an eventual quantitative formulation.

Visual Adaptation for Color

The sensitivity of the eye to light varies over a tremendous range, depending on the conditions under which it is used. It increases in dim light and decreases in bright in such a way as to tend toward a constant effective response. This process is known as brightness adaptation. Maximum sensitivity is obtained after long periods in total darkness.

A similar, although less well-known, phenomenon occurs when the eye is exposed to colored light, the color sensitivity decreasing in such a way that the effectiveness of the light as color is lost in large measure. This is known as color adaptation. The fundamental problem which must be solved before it is possible to lay down rules for color photography concerns the appearance of given color stimuli when the eye has undergone such an adaptation. If a photograph of a scene viewed and photographed by artificial light is to be viewed by either daylight or by artificial light, it is necessary not only that the photograph match the subject, point for point, for color, but also that if there should be any change in the appearance of the scene under the two conditions the photograph must undergo a similar change. It is also necessary that any mutual effects of adjacent areas be the same in the two cases and that they should change with the illuminant color in the same manner in the two cases. Obviously both the characteristics of the eye and those of the particular absorbing materials will play a part.

If the eye is rested for some time in total darkness, any effect due to previous stimulation by light wears off and the eye attains not only its maximum sensitivity but also what may be called its *natural sensitivity to color*. For the eye in this rested state there exists a certain range of colors which, when presented without previous eye stimulation, are accepted as achromatic. The *exact* quality of the color most frequently chosen is not known but there is good general agreement even though no international standard of quality has been set up. It approximates closely to what may be called noon sunlight. As a matter of convenience, it may also be considered equivalent either to the sun outside the earth's atmosphere (having approximately the same quality as the standard I.C.I. illuminant *C*) or to the stimulus corresponding to the light from an equal-energy spectrum. The sensitivity relations of the color receptor system of the eye which produce an achromatic response when exposed to this color may be thought of as the normal or resting state of the eye. The color itself may be called "absolute white."

It is not necessary here to postulate a definite receptor system for the eye nor to consider the number or nature of receptors involved. It has

[7] W. D. Wright, Proc. Roy. Soc. **115B**, 49 (1934); Ibid. **122B**, 220 (1937); Brit. J. Ophthalmol. **23**, 51 (1937).

[8] J. F. Schouten, Dissertation, Utrecht, 1937; J. Opt. Soc. Am. **29**, 168 (1939).

been adequately demonstrated that the sensitivity of the eye can be represented by sensitivity distribution functions for three assumed primary sources. Such a set of distribution functions has been standardized in the I.C.I. system of colorimetry. It will be assumed in what follows that when the response of the adapted receptor system can be represented by the same ratios of stimulus-sensitivity products as those calling forth white in the rested eye, the perception will be that of white or gray. A similar assumption is made with respect to stimuli which produce chromatic perceptions, i.e., the color perception is determined by the response due to the products of the sensitivities and the stimulus, the sensitivities being those of the three assumed receptors and the responses being modified by adaptation. For this purpose it will be necessary to distinguish clearly between the sensitivity of the eye to color in the sense of the relative amounts of energy at two different wave-lengths required to produce the same response from a given receptor and the sensitivity of the eye in the sense of the total amount of light necessary to produce the same response under two differing conditions. For the former the word "sensitivity" will be used, the latter, for want of a better term, will be called "responsiveness" but without any implications as to *speed* of response. That is, "sensitivity" will refer to the manner in which the eye integrates the energy distribution of a stimulus to produce a response, and "responsiveness" will refer to the magnitude of that response with respect to the integrated energy.

Wright[7] investigated the adaptation process for the fovea by means of a binocular matching technique. In this work he made use of the fact, which he demonstrated, that one eye may be completely adapted to a given stimulus without interfering with the state of adaptation of the other. By viewing fields with each eye so that they are seen as adjacent, he was able to make direct comparisons between the stimuli necessary to produce identical perceptions from each eye. In brief, he found that the energy necessary to evoke the same perception in the adapted eye as in the unadapted eye is directly proportional to the energy of the light used for adaptation.

If A is the intensity of the adapting field, a is the (very much smaller) intensity of the field as viewed by the adapted eye, and a_0 is the intensity of the field which must be presented to the other eye to produce a match at zero time after removal of the adapting stimulus, then $A = a/a_0$ or $Aa_0 = a = \text{constant}$. In his results, this relation is independent of the value of A. The equation states therefore that the responsiveness of the adapted eye is inversely proportional to the stimulus causing the adaptation. Wright found that this relation also holds for chromatic adapting stimuli and he demonstrated that the responsiveness to each of the primaries of a three-component mixture was reduced in direct proportion to the amounts of the primaries represented by the adapting light.

Wright was careful to point out that these relations had been examined only for the fovea and that they probably did not hold outside of the region over which the Fechner fraction was reasonably constant. He also found that some discrepancies exist in the blue region. He was concerned primarily with the rate of recovery of the responsiveness after adaptation and this he showed takes place linearly and fairly rapidly. The actual rates of recovery depend sharply on the intensity of the stimulus, and at high intensities the recovery ceases to be linear except in the initial stages. His work was done almost wholly with adaptation times of three minutes.

Schouten,[8] using essentially the same technique, investigated the effect on responsiveness of light sources in the field of view but not falling on the region of the retina being tested. He also carefully investigated the rate of decline as well as recovery of responsiveness as determined by both the time and intensity of the adapting exposure.

Schouten's results confirmed those of Wright except that his data do not show quite as constant a product of A and a_0. For adapting radiation not falling on the region tested, he found exactly the same effect but less marked, that is, adaptation was found in adjacent regions of the retina but the extent of the adaptation was less than in that of the exposed region. The actual amount found depended on the distance along the retina and on the intensity.

The rate of recovery of the adapted region was found to depend strongly on the time of exposure as well as on its intensity. Brief exposures caused

an adaptation from which recovery was extremely rapid and long times or higher intensities caused proportionally longer recovery times.

The rate of loss of responsiveness was found to be extremely rapid under all conditions, being essentially complete in a time of the order of 0.2 second. This holds equally well for adaptation in adjacent regions and in those directly stimulated.

No adaptations were found which applied to the eye as a whole. In all cases, the local areas of the retina, except insofar as they were near an exposed region, retained their normal sensitivities.

From these considerations and their abundant confirmations in the results of other workers (especially in the field of glare), a clear picture of the process of color adaptation can be assembled.

When an observer looks at a scene, his eyes are in almost constant motion. Detailed vision, however, occurs only when the eyes hesitate in a particular position. Whenever this occurs, adaptation takes place. Transference of the gaze produces a readjustment and this goes on throughout the scene. The speed of this readjustment, however, depends on the length of time that the gaze was held steady. Several times a minute, the eyes are closed for a fraction of a second.

From these facts it is seen that there are three types of adaptation which always exist simultaneously in the eye. The scene as a whole consists of brightnesses that lie between fairly well-defined limits which, in terms of the possible range for the eye, are not unduly separated. The eye accordingly does not receive stimuli outside this range for long periods at a time. Hence, there is an adaptation for the eye as a whole, controlled by something like the average intensity for the scene as a whole, from which recovery is quite slow. This may be called *general adaptation*. It does not exist in certain special cases, such as protracted viewing of a small isolated field. Superimposed on this are the adaptations due to the immediately effective stimuli and from which recovery is almost immediate when the gaze is transferred. These may be called *local adaptations* and are not noticed ordinarily unless the eye hesitates longer than usual on a particularly bright stimulus. Coincident with these is the "sideways" effect of local areas on each other, a brighter area decreasing the responsiveness of the eye to an adjacent area. This effect may be called *lateral adaptation*.

Each of these types of adaptation produces sufficiently different effects so that they are known in the literature as separate phenomena and they will be discussed separately under appropriate headings. All three play very important roles in color photography as well as in everyday life.

General Color Adaptation—Color Constancy

In all cases in which the eye responds to color, it may be considered as a mechanism which always integrates the spectral energy distribution of the stimulus. So far as the resulting color perception is concerned, the actual energy distribution is not a factor but only the integrals of this distribution with respect to the eye sensitivity system. The work of Wright and Schouten has shown that it is not these sensitivities which are affected by adaptation but what has been called earlier in the paper the responsiveness of the color receptors. In terms of these receptors, it is not the relative sensitivities as a function of the wave-length which have changed but the relative *outputs* of the receptors. This work shows further that color adaptation is the equivalent of a *division* of the responsiveness or output of the receptors by the *integrals* of the adapting stimulus with respect to these receptors.

The facts so far presented may be set up formally with respect to a three-color primary system. It will then be found that some modifications are necessary to bring them into exact accord with experience but a valuable technique is obtained by means of which complex viewing situations can be analyzed.

When the light from an illuminant falls on a selectively reflecting surface and is received by the eye, the energy distribution at the eye may be calculated by multiplying the incident energy of the illuminant wave-length for wave-length by the reflectance of the surface. The effectiveness of this light as color can now be determined by multiplying the curve separately by the three mixture curves of the system and determining the integrals under each of these product curves. The values of these integrals specify the amounts

of each of the corresponding primaries necessary to produce a visual match. The ratios of these values to the sum of the three, uniquely determine the degree of divergence of the stimulus from any fixed reference point.

In the I. C. I. system as customarily used, "dominant wave-length" is defined as the wave-length at which the line connecting the illuminant point (at least when the illuminant is Illuminant *C*) with the calculated point cuts the spectrum locus, thereby defining a direction of divergence. The ratio of the distance along this line, from illuminant point to stimulus point, to the distance from illuminant point to the spectrum locus is defined as excitation purity and establishes the amount of the divergence. Similar concepts are used by Judd[9] for his uniform chromaticity scale triangle.

Davis[10] has used a term "conjunctive wave-length" which defines a spectral wave-length with which any source may be mixed to obtain a stimulus corresponding to any point on the line connecting the two.

It is apparent from the definitions and assumptions on which these terms are based that they deal purely with the psychophysical aspects of the stimulus and are not intended in any way to describe or define the psychical color perception which will result from viewing the stimulus. Two stimuli having identical dominant wave-lengths and excitation purities will match if viewed under identical circumstances. But this does not state what their hue, saturation, and brightness (or lightness) will appear to the eye.

The work briefly outlined, however, suggests the possibility of a first approximation method of calculating the perceived color of a stimulus by taking into account the condition of the eye at the time of viewing.[11] From such calculations the mode of functioning of the eye can be made apparent, and calculations can be made of the probable change in *appearance* of a stimulus when it is seen under *different* eye conditions. This leads to a direct determination of the condition for "color constancy" of stimuli (again to a first approximation only, since psychological as well as psychophysical factors are involved in many complex viewing situations).

Wright's work indicates that the relative responses of the eye receptors under a given set of conditions can be obtained by dividing the integrals of the stimulus with respect to each receptor, by the integral of the adapting luminant with respect to the same receptor. On careful consideration, however, it is apparent that there is only one possible set of sensitivity distributions for three primary receptors for which this result will hold. Wright[1] determined such a set and referred to them as the "fundamental" sensitivities of the eye. Walters[12] has since redetermined them in the red and green regions and has thrown some doubt on Wright's values in the blue. The possibility of determining the set has, however, been demonstrated as has their similarity to the "grundempfindungen" curves of König and Dieterici.[13]

Let X, Y, and Z represent the integrals under the curves obtained by multiplying the red, green, and blue receptor sensitivities, respectively, by the energy distribution of the stimulus. Then, let X_{SD}, Y_{SD}, and Z_{SD}, be these integrals for a sample S and illuminant D, and X_D, Y_D, and Z_D be these integrals for the adapting illuminant (also D). Under the assumptions, these values are linear functions of the receptor outputs; the relative outputs then become X_{SD}/X_D, Y_{SD}/Y_D, and Z_{SD}/Z_D.

If a Maxwell triangle be constructed based on X, Y, and Z for "white" light (conveniently the equal energy spectrum), then these values determine a point for any stimulus and adaptation condition (illuminant quality being taken the same as the adapting radiation). This point is calculated in the usual manner:

$$x_{SD/D}=\frac{X_{SD}/X_D}{X_{SD}/X_D+Y_{SD}/Y_D+Z_{SD}/Z_D}$$

[9] D. B. Judd, J. Opt. Soc. Am. **25**, 24 (1935).

[10] R. Davis, Bur. Stand. J. Research **7**, 659 (1931).

[11] Although the present writer was led to this concept chiefly by the work of Wright, it was suggested in part, or in whole, by several previous writers. So far as he can determine, it was first proposed by H. E. Ives as a pure hypothesis to explain his observations on artificial daylight: H. E. Ives, Trans. Illum. Eng. Soc. **7**, 62 (1912). The following references to the literature using the same approach have been kindly supplied by Dr. Judd: E. Noteboom, Zeits. f. Instrumentenk. **55**, 317 (1935); W. Ströble, Das Licht **9**, 149 (1939).

[12] H. V. Walters, Proc. Roy. Soc. **131B**, 27 (1942).

[13] A. König and C. Dieterici, Zeits. d. Psychol. u. Physiol. d. Sinnesorgane **4**, 241 (1892).

and similarly for $y_{SD/D}$ and $z_{SD/D}$.

$$x_{SD/D}+y_{SD/D}+z_{SD/D}=1.$$

For this Maxwell triangle, the illuminant point is always at the center for all illuminants, but a different spectrum locus is obtained for each illuminant.

A line from the center through a stimulus point cuts the spectrum *locus for its illuminant* at a point which has the same significance as the "conjunctive" wave-length of Davis. Purity, if desired, could be defined as usual by the relative distance of the point and its spectral locus. In these respects, the triangle so constructed is no different from any other possible triangle. The possible advantage arises from the fact that *under the assumptions* each point of the triangle is *invariant with illuminant*. That is to say, if two illuminant-sample combinations calculate to the same point, they will be perceived as identical when each is seen under its own illuminant. This property is not possessed by a triangle based on any other primaries and holds for this one, of course, because of the method of calculation.

It accordingly becomes possible, by means of this construction, to tell whether or not a given sample will appear as the same or different under two different illuminants. The sample is simply calculated for both illuminants, and if the points superimpose, the sample will appear the same; if they do not, it will appear different.

Having gone this far, it may not be out of place to point out some other possibilities inherent in this construction, but it must be kept in mind that the method described is inherently a first approximation. For example, the invariance of the illuminant point indicates that under all conditions, when the eye is adapted to an illuminant, it will appear achromatic. This is known not to be true but it does hold to a first approximation for continuous sources of an energy distribution not too far from that of a blackbody (daylight and tungsten light, for example).[14]

It also is in error within a given illuminated area if there is any great intensity range involved, since under these conditions Helson[15] has shown that the brightest achromatic objects appear tinged with the color of the source (as seen under approximate daylight adaptation) and the approximate complementary of this in the dimly lighted regions. These effects, however, are large only for strongly chromatic sources, so that again a fair degree of approximation is indicated.

With these limitations in mind, it is interesting to see how the result could be interpreted if a sample in two illuminants gave points which did not superimpose. The problem in this case is to determine the proper interpretation of the direction and magnitude of the divergence. Neither the constant hue nor the constant saturation lines will be straight in such a triangle and these attributes in general will shift with intensity. However, since every point corresponds to a given perception, each point may be correlated with a known standard as seen under known conditions. A very suitable set for this purpose would be the Munsell system as seen under daylight conditions against a background of a given reflectance. This set has already been correlated[16] with the I. C. I. system, and it would only be necessary to obtain the conversion equations for the two systems. It would then be possible to define directly the shift in appearance of the color by *specifying the two appearances* in the Munsell system.

A further interesting point is the appearance of the monochromatic stimuli under various adaptation conditions. The appearance of these, relative to daylight, shows directly on the triangle, since the plots of the loci for different adaptations can be compared directly with that for daylight. Purities higher than unity are readily obtained in some regions, such as the green, while other points, such as the extreme red, are invariant.

As has been pointed out repeatedly, this whole construction is an approximation, and as such, it must be compared with others experimentally before it can be determined whether or not it is the best one. This is true whether or not it does portray the mode of action of the eye receptors correctly.

---

[14] L. T. Troland, J. Exp. Psychol. **4**, 344 (1921); H. Helson and D. B. Judd, *ibid.* **15**, 380 (1932).

[15] H. Helson, J. Exp. Psychol. **23**, 439 (1938); H. Helson and V. Jeffers, *ibid.* **26**, 1 (1940).

[16] See, for example, the Symposium on the Munsell system in J. Opt. Soc. Am. **30**, December (1940); especially J. J. Glenn and J. T. Killian, "Trichromatic analysis of the *Munsell Book of Color*" and K. L. Kelly, K. S. Gibson, and D. Nickerson, "Tristimulus specification of the *Munsell Book of Color* from spectrophotometric measurements."

Judd has pointed out,* for example, that the work of Adams[17] on the x–z coordinate plane (I. C. I. system) has given promise of accurate correlation with the known facts, and it may well be that this will be found to be a better approximation. They cannot be the same since different receptor systems are assumed, and if Wright's work is correct in the main, the x–z system cannot correctly portray the action of the receptor system.

The postulated mode of action of the eye and the use of the fundamental primaries permit a mathematical statement of the conditions for complete "color constancy" of a sample-illuminant pair.

Using the same nomenclature, but introducing two illuminants D and D' and considering the x value only (exactly similar considerations hold for y and z), the "condition" may be set up formally as follows:

$$x_{SD/D}=x_{SD'/D'}, \quad y_{SD/D}=y_{SD'/D'}$$

and

$$z_{SD/D}=z_{SD'/D'},$$

$$x_{SD/D}=\frac{X_{SD}/X_D}{X_{SD}/X_D+Y_{SD}/Y_D+Z_{SD}/Z_D}$$

and

$$x_{SD'/D'}=\frac{X_{SD'}/X_{D'}}{X_{SD'}/X_{D'}+Y_{YSD'}/Y_{D'}+Z_{SD'}/Z_{D'}}.$$

It is seen that these expressions both reduce to x_S (i.e., the value for the equal energy system here defined as "white"), if $X_{SD}=X_S\cdot X_D$ and $X_{SD'}=X_S\cdot X_{D'}$. In other words, if the product of the *integrals* of the distributions integrated separately is the same as the integral *after* the products have been performed, then the color *appears* the same under both illuminants. This condition is obviously met by definition for all nonselective surfaces. For all other surfaces, the surface absorption distribution and the illuminant energy distribution play equally important parts.

For the most important case (D = daylight and D' = some continuous distribution which is not greatly different), the degree of approximation of this equality is quite good for most ordinary reflecting surfaces since most surfaces have broad, gradually sloping absorption curves. For highly selective surfaces or for discontinuous or sharply selective light sources, the approximation may become very poor indeed, leading to a distinct change from the daylight color. The generally good approximation of most surfaces leads to the concept of "approximate color constancy" in everyday life. The poor approximation under certain illuminants gives rise to the large shifts of perceived color seen, for example, under many of the commercial fluorescent light sources.

Another interesting situation arises when the adapting illuminant is a mixture of two light sources of different chromaticities. Regardless of the chromaticities, the mixture will be represented by the achromatic point at the center of the triangle, and the points representing the two sources will be collinear but on opposite sides of the point. Since the points on this triangle may be interpreted as representing the appearances of the stimuli, and since collinear points on opposite sides of the center are complementary in daylight, it follows that, regardless of the nature of the two sources, they will appear in the customary hues of complementary colors. This is the well-known phenomenon of complementary shadows.[18] The effect is so powerful that two-color additive pictures can be made to produce a satisfying greenish blue by projecting one picture through a yellow and the other through a red filter. In this case, the yellow becomes the blue to the adapted observer and the red becomes a complementary orange.*

A similar consideration holds also for the calculation of local and lateral adaptation effects. The results are obviously similar to these and the effects are well known. Local adaptation to a chromatic surface followed by viewing of an achromatic one leads to a complementary "after image." If the second surface is also chromatic, the result is an apparent "mixture" with the complementary. These effects, of course, are familiar as "successive color contrast" phenomena. Lateral effects produce the same result.

* Private communication.

[17] E. Q. Adams, J. Opt. Soc. Am. **32**, 168 (1942).

[18] See, for example, J. Plateau, *Bibliographie analytique des principaux phenomènes subjectifs de la vision jusqu'à la fin du xviii^e siècle* (Royal Academy of Science, Bruxelles, 1877) and more recently, J. P. C. Southall, *Physiological Optics* (Oxford University Press, New York, 1937), p. 398.

* Demonstrated by the writer before the Optical Society of America, March 5, 1943 in New York.

Adaptation of a local area by a chromatic surface leads to a decreased responsiveness of the adjacent areas. If these areas are receiving light from a second surface, the perceived hue is shifted. These effects are mutual and give rise to the phenomena of "simultaneous color contrast." These effects will be considered in more detail in a later section.

In general, in viewing a scene, first one color and then another affects the eye as it moves about. All the colors will be modified by the illuminant. For most of them, color constancy will be good. The general adaptation *condition* of the eye, therefore, will approximate that of the light source at an intensity level considerably below that of the source itself. For some special scenes, this will no longer be true. Consider from this standpoint a simple two-part field, in which a small area is seen surrounded by a much larger one which nearly fills the retinal area. If the central area is colored and the larger one is achromatic but the whole field is illuminated by light of non-daylight quality, two extreme cases arise. If the background is black and the eye moves freely about, there will be very little general adaptation and the eye will remain in its normal color state. The patch in this case will have the same appearance as a surface in daylight having a reflectance distribution equal to the product of the reflectance of the surface and the energy distribution of the actual illuminant. That is, it will be seen as a stimulus having the energy distribution of the product curve of the surface and the source. If the background is white, however, the eye becomes generally adapted to the color of the illuminant and the central patch now looks the same as it would in daylight, so far as the color constancy properties of the surface and illuminant permit.

If both the central and the surround patches are colored, the adaptation approximates that caused by the surround. As just shown, this causes the effective stimulus from the small central patch to act as though its product integrals had been divided by those of the surround. This is the customary movement of a color in the direction of the complementary of the surround which is seen in all simultaneous contrast effects.

These points are mentioned to distinguish the effect of general adaptation from that of lateral adaptation or simultaneous contrast to be discussed in the next section.

Lateral Color Adaptation—Simultaneous Color Contrast

There is considerable confusion in the literature concerning the change in the appearance of one color when another is placed adjacent to it. Since almost all color vision involves this phenomenon, it is extremely important that a unified approach be available. These considerations on general adaptation plus the work of Schouten mentioned in the first section make such an approach possible. They do little, however, to explain or define the almost indeterminate nature of any complex viewing situation except to make the effects understandable.

Schouten showed that adaptation of any local area produced a corresponding adaptation of adjacent regions, the magnitude of the effect decreasing with the distance from the exposed area. The responsiveness of the receptors is accordingly depressed nearly as much, immediately adjacent to as in the region of the exposure. The effect, however, decreases fairly rapidly as the distance from this area increases. The effect at any given distance is roughly linear with the intensity of the light in the exposed area. Recovery, as before, depends on both the time and the intensity of the exposure.

There are two general effects to be observed from the point of view of color perceptions. First, because this lateral effect tends to spread out the effectiveness of a bright spot of light over the retina, it combines with the eye movements to produce a more uniform general adaptation. Second, the depression of responsiveness in adjacent areas modifies the perception of the chromaticity of an adjacent stimulus. If two areas are of comparable size and brightness, this effect is mutual. If one is more saturated than the other, or particularly, if it surrounds a smaller area, the effect is seen primarily in the smaller or less saturated of the two.

It will be noted that this effect is exactly the same in many cases as the effect of general adaptation in a two-part field, i.e., adjacent colors tend to move each toward a mixture with

the complementary of the other. The lateral effect, however, is purely local and acts quite independently of that due to general exposure. If an achromatic area is seen surrounded by a color, there is a strong tendency for the achromatic part to take on the color of the complementary to the surround and this is independent of colors in other parts of the field. Even in the simple two-part field consisting of a small patch surrounded by a much larger one, lateral adaptation may be demonstrated. Constant eye motion from point to point over the whole field will show a very strong adaptation effect on the central field while fixation of the eye on the central spot will be found to show a much diminished effect which frequently falls off toward the center of the area. Involuntary slight displacements of the fixation point will reveal sharp-edged adaptation regions directly underlying the image of the surround. Under these conditions the diminished induced color of the central area is due to lateral adaptation while the bright edges are due to the local adaptation which will be discussed in the next section.

Lateral effects may obviously be calculated by the same technique used for general adaptation as was pointed out earlier. In this case, the integrals by which the responsiveness of each receptor is divided must be modified to allow for the effects of distance, intensity, etc. Since it has been shown that this approach in itself is qualitative, it is hardly worth while to set up the equations by which such effects may be calculated. They follow the same procedure as was considered for general adaptation but their validity is limited to the *instantaneous* condition existing after each eye movement. That a colored surround about an achromatic area produces a complementary color sensation in that area follows from reasoning similar to that used for shadows illuminated by two colored sources. If the responsiveness of the eye to green is decreased, an equal energy stimulus will cause a greater response from the blue and red and hence will look magenta, complementary to the green.

This lateral type of adaptation is important to color photography in two ways, first, as it affects a print as a whole, owing to comparison with the surroundings, and second, as a relative area function within the print. Little exact information seems to be available about this latter function as far as color is concerned. Two definite qualitative effects are known but require investigation before they can be stated quantitatively. The effectiveness of a given area of color is greater as such, i.e., the color appears more intense the greater its area. This is probably due to a diminished brightness contrast effect from the surround in most cases and will be discussed in a subsequent article. From the work of Schouten, the smaller the area the more it is affected by the surrounding colors. The combination of these two effects implies that any exact color reproduction of a scene must meet the requirement that the individual parts subtend the same angle at the eye as the original scene or *be suitably modified in accordance with these effects* in order to give a convincing likeness. No single law has yet been deduced for such a modification but practical experience has indicated the necessity that it be done.

It is important to note that simultaneous color contrast effects can be observed equally well when the intensity of the adapting stimulus is so low that it does not, by itself, produce a chromatic perception. A saturated but dark green surround will make a small central patch of gray look magenta at intensity levels which are so low that the surround appears black.[19] It is because of this and similar facts that a separation of brightness and color adaptation is justified.

Although little or no work seems to have been done in this field, it must be true that for relatively large visual angles simultaneous contrast effects vary considerably with the particular area which is being fixated, with the length of time the area is fixated, and perhaps even with the particular aspect of the object to which the attention is directed. It is certainly true for brightness considerations and it seems likely to be true for color in general. A very interesting feature of all simultaneous color contrast work and one which deserves more attention than it has received is the fact that the amount of "induced color" is partly under voluntary control. In this aspect, the subject passes over into the field of pure psychology. One example will be given to make

[19] W. von Bezold, *The theory of color* (L. Prang and Co., Boston, 1876), p. 163, S. R. Köhler, translator.

clear the type of phenomenon discussed, although many others could be cited.

If a series of sheets of gray paper are interleaved with green in such a way that only a small strip of the gray sheets is visible, the following situations can be verified easily. If the pile of sheets is laid on a large black background, the gray sheets appear a strong magenta. If the pile is placed on a large gray sheet, this tendency will be much decreased, especially if one of the gray sheets is on the bottom. Even on the black background, however, the effect can sometimes be made to disappear completely by simply laying a narrow gray strip across the pile so that it is available for direct comparison with all the strips. This effect, in common with many others (such as the fact that a coin seldom shows any trace of color contrast regardless of its background), may be summarized by the statement that color contrast may be inhibited if the actual situation is sufficiently obvious.

Local Color Adaptation—Successive Color Contrast

The remaining type of effect produced by color adaptation is that due to momentary viewing of the scene. Since all adaptation is local, the phenomena of this phase of the subject are distinguished almost entirely by the rapidity with which the eye recovers its responsiveness. Fixation of the eye on a particular area for a brief time, followed by transference of the gaze to another surface, gives rise to characteristic "after-images" of the first surface. The phenomenon as a whole is known as *successive color contrast*. The color sensation produced is indicated qualitatively by considering the second surface from the standpoint of receptors whose responsiveness has been modified by the integrals of the first surface. These effects wear off very rapidly and not at equal velocity for the different receptors, giving rise to a phenomenon known as the "flight of colors."

Such effects are not ordinarily observed in daily life, partly because they are transient but more perhaps because the blinking of the eyelids permits time for almost complete recovery from all but the strongest effects.

Successive color contrast again gives evidence of the independence of brightness adaptation and color adaptation. A colored stimulus whose intensity is too low to cause the sensation of color will produce a complementary colored after-image when the gaze is transferred to a brighter achromatic surface.[19] This fact, incidentally, seems significant in connection with some forms of color-blindness. Green-blind people are sometimes able to see a brilliant red after-image following stimulation by green light which to them appears gray. (This has been found to be true by MacAdam* in the case of an almost complete dichromat, a deuteranope.) This implies that the cause may be a loss of responsiveness of the green receptor for chromaticity while retaining its responsiveness to brightness. It is, perhaps, the responsiveness of the green receptor system to color relative to its responsiveness to brightness, rather than its color-sensitivity distribution as such, which is abnormal.

Indeterminacy of Color Perception

In the preceding sections an attempt has been made to clarify the bases on which predictions may be made about the appearance of a given color stimulus. It is apparent that for a given material with a given spectral reflectance, in surroundings whose characteristics are completely known, a good estimate may be made as to its appearance, especially with regard to the effect of a change in the illuminant. In the actual perception of color, however, little actual determinacy is apparent. If an observer is given an object with a dark desaturated color in artificial light and asked to describe its probable appearance in daylight, he will have considerable difficulty in reaching a decision. He will ultimately place it in the strongest light available, perhaps next to a white, and then guess. The reasons for this are evident. Few objects show exact color constancy and the exact adaptation condition for the eye depends on its recent history. The brightest possible light and direct comparison with white are the best conditions attainable.

The fluctuation of apparent color with eye movement is well illustrated by small color patches mounted on large white and black backgrounds. This observation has been carefully

* Private communication.

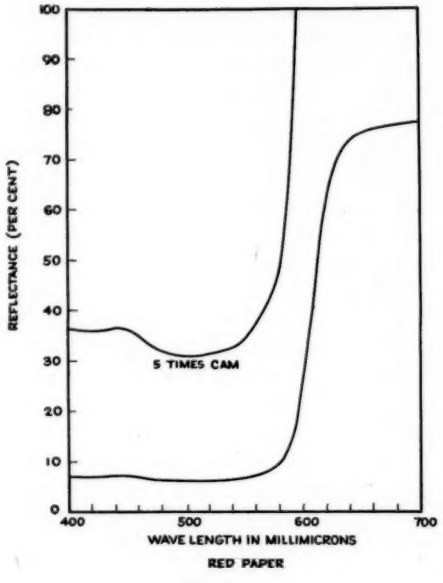


FIG. 2.

considered by Judd.[20] Because it offers a nice check of the theory that absolute white represents the natural ratios for receptor responsiveness, an interesting experiment was tried. A colored paper showing poor color constancy was chosen. The spectrophotometric curve is shown in Fig. 2. A 2-inch square of the paper was mounted in the center of an 11- by 14-inch white card and another on a similar black card. In artificial light, these two showed a strong tendency to fluctuate in appearance as one looked from one to the other while holding the cards in the hand. The piece on white was usually seen as a slightly purplish red and showed the least variation, whereas the one on black varied from a distinct orange through orange-red to an occasional instant during which it appeared to match the one on the white. The causes of this fluctuation have already been discussed in terms of the changing adaptation of the eye. It should follow, however, that if this explanation is true there should be no such chromatic fluctuation when the cards are viewed in light of quality approaching that of absolute white. This was found to be the case. In good daylight, no fluctuation was observed, the two patches consistently appearing to match for color. It is again interesting to note that most observers reported they could see no difference in the relative lightnesses of the two under the two conditions, although under both conditions some lightness fluctuations were observed, and of course the one on black appeared the lighter of the two.

This indeterminacy of color vision does not cause us concern; in everyday life we are accustomed to thinking of most colors as not changing at all. This is in large part due to the tendency to remember colors rather than to look at them closely. For the most part, careful observation of stimuli is made only by trained observers. This same tendency also operates favorably in viewing color photographs. It is seldom necessary to obtain exact color reproduction of a scene to obtain a satisfying picture. It is necessary, however, that the reproduction shall not violate the principle that the scene *could* have thus appeared.

The Problem of Chromaticity Reproduction

The basic first-order problem of color reproduction may now be restated briefly. The *effective* chromaticities of the subject must be reproduced by areas which have as nearly as possible identical *effective* chromaticities. For scenes in daylight for which the reproduction is to be viewed by daylight, only the usual chromaticities with respect to the *C* point are involved insofar as eye adaptation phenomena are concerned. If the print is also to look satisfactory in artificial light, the dye system of the process must show maximum color constancy. This sets a severe limitation on the sharpness of cut of the dye absorptions which may be used and this, in turn, tends to limit the saturation of the colors in the reproduction. The constancy requirement is especially severe in respect to the absorption curve of the mixture which represents gray. For the general case of a process which is to be viewed under *any* illuminant, the dyes must mix to a neutral which is *spectrally nonselective* in order that the grays will not change color with that of the viewing light.

If a scene is to be photographed by artificial light so that the reproduction will look as though

[20] D. B. Judd, J. Opt. Soc. Am. **30**, 2 (1940); Nat. Bur. Stand. J. Research **24**, 293 (1940).

it had been photographed by daylight, an additional set of requirements are added, not all of which can be met by a three-color system. Achromatic surfaces in the subject must be reproduced as nearly by achromatic deposits as the process permits. In a three-color process, each color being controlled by a separate photographic emulsion, each emulsion has the same pairs of sensitivities as those already considered for the visual mechanism. Each has a wave-length distribution of relative sensitivity and each has a responsiveness to energy ordinarily called the "speed" of the emulsion. In order that an achromatic surface illuminated by artificial light may give a neutral deposit, the relative values of responsiveness or speed of the emulsions must be adjusted so that the *effective response* is the same to artificial light as would be required to give the same deposit by daylight. The situation is exactly analogous to that in visual color adaptation and produces the same type of result. For a colored surface, therefore, the integrals of the product curve of the light source and the surface are divided by the integrals of the light source itself to specify the ratios of dyes obtained. The integrals in this case are those with respect to the spectral sensitivities of the film rather than the eye. (They would produce the same result if the film had the "fundamental" sensitivity distributions of the eye receptors.) Surfaces which, in general, show good visual color constancy will normally be fairly well reproduced. Colors showing poor color constancy, however, will not only photograph differently than they would appear in daylight but also differently than they appeared in the artificial light. For these colors, therefore, reproduction will be poor. This is, however, the only *type* of solution possible. The widespread desire to photograph scenes by artificial light makes it necessary for manufacturers to supply two kinds of film "balanced" for the two types of light source. The fact that most surface colors show good constancy plus the various factors discussed later which frequently aid visual acceptance makes such a practice almost entirely satisfactory. An occasional attempt to reproduce a brilliantly colored fabric or flower may be a complete failure in the same process. Viewing the reproduction by artificial light, of course, does not correct the situation since the color constancy failure of the dye system is entirely independent of that of the original subject.

One other case may be considered here. If a scene is to be photographed in artificial light so that the impression of artificial light is given when the reproduction is viewed by daylight, a serious dilemma is encountered. Under artificial light, as pointed out, the source itself does not lose all of its visual color. Artificially lighted rooms (using incandescent bulbs) do not appear to be illuminated by white light but by a distinctly more yellow light than daylight. This may be due to a nonlinearity in the adaptation mechanism, but this same sort of effect could be introduced into the film since the linearity of response with exposure is more or less under the control of the manufacturer. Such film, however, could not also be used to represent effective daylight and would have no exposure latitude whatever. For these reasons, and because it is usually possible to simulate artificial light by means of lamps, etc., in the picture, such films are not manufactured. Where necessary, lamps of lower color temperature than that for which the film is balanced can be used for local lighting to suggest the required effect still further.

This discussion indicates the most serious single limitation of present-day color photography. As the observer shifts from scene to scene and even from point to point in one scene, the color adaptation of his eyes is constantly shifting. He is aware of this only if his attention is called to it and even then can observe it only because of the slight residual source color which the eye does not eliminate, or because of memory of the daylight appearance of some color which shows poor constancy. This is not true of the photographic film with which he takes pictures. Two types of film are available but each of these corresponds to one and only one color adaptation condition of the eye. *Unless the observer's eye adaptation condition corresponds to the illuminant for which the film which he is using is balanced, the reproduction will not correspond at all to what he sees.*

The points which have been raised in this section are independent of the relation between the sensitivity distributions of the emulsions used and those of the eye. For the sake of completeness it might be well to point out that in addition to

the above effects unless the emulsion sensitivities are valid transformations of eye mixture data with respect to the dye primaries of the process, practically no point for point chromaticity matches will be obtained under any conditions except perhaps in the neighborhood of gray. Chromaticity matches in the subject do not, in general, exactly match in the reproduction in any known photographic process. It is doubtful if such a process is possible.

If the equivalent of color constancy shifts are to be obtained by changing the speeds of the emulsions, then the emulsion sensitivities must be identical with those of the eye.*

Commercially available color processes, however, have been developed empirically to give the best possible compromise under the conditions to be encountered. The success of these compromises measures the commercial success of the particular process. As a matter of fact, the errors due to this cause are far less serious than those caused by the frequent failure of the photographer to use illumination of the correct color.

Reproduction of Simultaneous Contrast Effects

The preceding section considered the factors involved in obtaining a reproduction in which the effective chromaticities of each area, *considered by itself*, matched the effective chromaticities of the corresponding area of the subject. When each area is considered also in relation to the adjoining areas, the relative lateral adaptation or simultaneous contrast effects must be taken into account. This problem is encountered when it is desired to make a number of reproductions of the same subject of different *sizes* so that they will all look alike, as well as looking like the original subject.

As pointed out before, only a few directly applicable data are available. Experience with color processes, however, has shown a number of essential requirements. Stated briefly, a small print must have both higher color saturation and higher contrast if it is to compare favorably with one five or six times the size. The required difference is so large, in fact, that if contrasts suitable for small pictures are used for large ones, the quality is inadequate.

At first sight, this is surprising since it is generally accepted that simultaneous contrast effects are greater on small areas than on large, and the intention is to produce the *same* effect as that produced by the subject. The increase with small areas, however, has been demonstrated only for cases in which the test area is reduced and the inducing area is retained at full size. The fact seems to be that when all the areas are reduced proportionately the effect is reversed. Painters of miniatures have also encountered the same phenomenon. In this type of painting, the maximum contrasts possible with pigments can be used with impunity in places which would ruin a larger picture.

Effect of Viewing Conditions on the Apparent Quality of Reproduction

Isolated Field Viewing

In the preceding section it was implied that a number of conditions under which reproductions may be viewed tend to aid the observer in seeing the reproduction as better than it actually is. The first and, perhaps, the most favorable of all viewing conditions is the familiar projection of a picture on a screen in a darkened room. In such a situation, the eye readily adjusts its adaptation conditions to that of the average of the picture on the screen. The result is roughly the equivalent of a correction in the taking of the picture to the most suitable light source for the material since the same *scene* also controlled the photographer's eye when he took the picture. The correction is not complete and a trained observer can judge the daylight appearance of a picture with fair success. The apparent magnitude of any errors present, however, is a small fraction of what would be seen if no shift of adaptation occurred.

Even a person who is familiar with this effect and with the enhancement of quality it produces is often surprised at the order of magnitude of the possible change. In terms of the color temperature of the projection lamp, a picture with a large viewing angle will appear acceptable at temperatures from below 2000°K to above 10,000°K, or roughly, say, from that of a kerosene lamp to

* It is interesting that the late F. E. Ives always insisted that the sensitivities of the three emulsions in a three-color process should exactly correspond to the fundamental sensitivity distribution curves of the eye receptors. The present analysis confirms this view for the general case.

that of a blue sky. Under any of these conditions, the observer is *aware* of the color of the light but the difference from daylight seems small and the pictures are entirely satisfactory. Observable distortions are actually more serious from the standpoint of relative brightness than from hue or saturation. At the blue end of the range, reds appear quite dark and, conversely, blues are quite dark at the yellow end.

The question is sometimes raised as to why it is necessary to have both a daylight and an artificial light type of film if such visual compensations can occur. As a matter of fact, such pictures can be taken but the latitude required of the color process is beyond the range of materials at present available. Such pictures are also useless unless projected in a *completely* dark room.

The subject has not been investigated thoroughly and acceptable variations are obviously a matter of personal taste. It is certainly true that relatively large errors in over-all color balance in any direction make little difference in the enjoyment of pictures under these conditions. Most observers are not even aware that such differences exist.

The nature and to some extent the magnitude of the phenomenon can be demonstrated by projecting a fairly wide border of light around the picture. This border will tend to control the condition of the eye and the pictures will then be seen under more or less constant conditions. Two phenomena are observed. Any appreciable deviation from correct color balance in the picture is immediately detected by any observer. If the color of the border is different from that of the projection light, all normal pictures look badly off in the complementary direction.

Successive Viewing

If one picture on the screen is quickly changed to another, the local adaptation caused by the first picture determines the appearance of the second. In many cases, e.g., in motion picture work, this limits sharply permissible variations in color balance. While it is true that the eye quickly adjusts itself to the new scene, the color changes may become so irritating as to spoil the picture for the observer. It is interesting that the requirement is lack of *difference* from one scene to the next. It is not equally important that the balance be correct for daylight viewing unless a surround is present. In the projection of slides, simply moving the slide carrier slowly rather than rapidly will frequently give enough time for the eye so that a large color difference will not be noticed.

Viewing in Relatively Dark Surroundings

An intermediate case between projection in a dark room and projection with an illuminated border is of considerable interest and importance. Large transparencies, intended to be viewed over an illuminator, are customarily seen under conditions in which the illuminator is by far the brightest area in the field of view. Under these conditions, and to the extent that this is so, the color constancy effects mentioned will correct improper color balance, provided the illuminator is completely covered with the transparency. If part of the bright area of the illuminator is uncovered, the picture will be seen more or less correctly with respect to this color. Hence, if the illuminator has roughly a blackbody energy distribution, its apparent color temperature does not matter over the range of 2400°K to 5400°K, a range smaller than that possible in dark-room projection. The extent of the compensations which occur, even under these apparently adverse conditions, is occasionally difficult to believe. If two large transparencies, for example, are deliberately so adjusted that one is correctly balanced and the other is distinctly green, and these are shown to an observer, one at a time, over the same illuminator, he may be completely unable to tell which one is correct. After the observer has viewed the green transparency, the other will at first appear quite magenta. Careful study will convince him that this one is quite good and the first will now look very green. This change of local adaptation is particularly noticeable if the two transparencies are now seen side by side. Under these conditions, adaptation changes to an *intermediate* point and *neither* of them looks acceptable. It may be stated as a general rule that if two pictures with even slightly different color balances are seen, side by side, an observer will state that the correct value lies "halfway between." This is frequently true even when both pictures are off balance in the same direction.

What has been said of transparencies applies also to large reflection prints, provided they are viewed in a sufficiently *isolated* position and are more brightly lighted than their surroundings.

Viewing in Normal Surroundings

If the reproduction is to be viewed while held in the hand, or is seen among equally illuminated and familiar objects, almost all these effects disappear insofar as they are beneficial to the picture. The observer then sees the print simply as one of the objects in the field of view and his eye is adapted to the general illumination. Under these conditions, a critical observer can and does note errors in the *over-all color* of a picture with a precision comparable to that which can be obtained in a two-part photometric field. Provided the "over-all" color or color balance is correct, the observer will be satisfied with the color reproduction of the *subject* if it meets one very interesting requirement. The process must be so adjusted that the colors are internally consistent. This "consistency principle" cannot be stated rigorously at the present time. Roughly, it requires that no hues which are familiar be badly off color and that the lightnesses and saturations of the colors each bear the correct relation to those of the subject. Neither saturation nor lightness need be equivalent to those of the subject but the rendition of some colors *must not be better* than that of others. In general, it is less desirable to have good reds and poor greens than to have both poor. If these conditions are met, the eye *sees the colors as closer approximations* to those of the scene depicted than they actually are. Borders around the picture enhance this apparent fidelity to the original by setting the print apart from its surroundings.

It is significant in this connection that the people who are daily engaged in working with a particular color process are perhaps the worst possible judges of its fidelity of reproduction rather than the best, as might be expected. This does not result from any feeling of partiality toward the process, as is sometimes imagined. Unconsciously, they learn to see the reproduction *as though it were* correct. Examiners checking the quality of prints in mass production have to pick up a print from time to time and move about the room or carry it to a window to prevent the occurrence of temporary sets in their judgment. This is especially true if all the prints in a given batch are running slightly off color in one direction.

The facts stated above have been obtained empirically by direct experience with color photography in large-scale production. The approach to the visual phenomena involved in color photography which has been used in the first part of the article was developed through an attempt to explain these photographic effects and bring them under control. Postulation of purely mental phenomena to explain apparent discrepancies between stimuli and perceptions has been avoided deliberately. It would be a mistake, in the writer's opinion, to assume that no such effects are present. It is hard to explain on purely physiological grounds, for example, why familiarity with a poor print may so greatly improve its appearance, regardless of viewing conditions. More information is needed, however, before such effects can be generalized constructively.

As stated at the beginning, brightness has been treated throughout this part as independent of chromaticity. Some justification for this has been given. Since a consideration of brightness involves the subject matter of the previous part as well as introducing much new material, the discussion of this phase of the subject will be presented separately.

Appendix to Part II

The colorimetric discussion in the central part of this article was originally written in terms of the I.C.I. Colorimetry system. Dr. MacAdam kindly pointed out to the writer that this system is not directly applicable, because only one system can meet the requirements laid down, and it is not likely that the I.C.I. primaries constitute such a system since they were chosen merely for convenience in computation. In support of this, he submitted the following demonstration.

Theorem:

If a stimulus matching a mixture of R, G, B units of a certain set of primaries has the same appearance to an observer adapted to R_A, G_A, B_A as does another stimulus R', G', B' to the same observer when adapted to $R_E=1$, $G_E=1$, $B_E=1$, and if $R'=R/R_A$, $G'=G/G_A$, $B'=B/B_A$ then this is the *only* set of primaries for which the tristimulus values of these adaptively equivalent stimuli are connected by equations of this form.

Proof:

Assume another set of primaries for which the tristimulus values are r, g, b. Then, in general,

$$r=a_{11}R+a_{12}G+a_{13}B$$
$$g=a_{21}R+a_{22}G+a_{23}B$$
$$b=a_{31}R+a_{32}G+a_{33}B.$$

Without loss of generality in this discussion:

$$a_{11}+a_{12}+a_{13}=1$$
$$a_{21}+a_{22}+a_{23}=1$$
$$a_{31}+a_{32}+a_{33}=1.$$

Then the first adaptation is to:

$$r_A=a_{11}R_A+a_{12}G_A+a_{13}B_A, \text{ etc., for } g_A \text{ and } b_A,$$

and the second adaptation is to:

$$r_E=1, \quad g_E=1, \quad b_E=1.$$

The second stimulus, R', G', B', is specified in the new set of primaries by:

$$r'=a_{11}R'+a_{12}G'+a_{13}B'$$
$$=a_{11}R/R_A+a_{12}G/G_A+a_{13}B/B_A, \text{ etc., for } g' \text{ and } b'.$$

But, by the use of a rule of the form stated in the theorem, the adaptively equivalent stimulus would appear to be

$$r''=r/r_A=(a_{11}R+a_{12}G+a_{13}B)/r_A, \text{ etc., for } g'' \text{ and } b''.$$

The discrepancies between the tristimulus values of the actual equivalent and that predicted by this misapplication of the rule are:

$$r''-r'=(1/r_A-1/R_A)a_{11}R+(1/r_A-1/G_A)a_{12}G+(1/r_A-1/B_A)a_{13}B$$
$$g''-g'=(1/g_A-1/R_A)a_{21}R+(1/g_A-1/G_A)a_{22}G+(1/g_A-1/B_A)a_{23}B$$
$$b''-b'=(1/b_A-1/R_A)a_{31}R+(1/b_A-1/G_A)a_{32}G+(1/b_A-1/B_A)a_{33}B.$$

Since r_A is in general not equal to G_A or B_A, the discrepancy $r''-r'$ can be zero only if $a_{12}=a_{13}=0$. Similarly, $g''-g'$ can be zero only if $a_{21}=a_{23}=0$ and $b''-b'$ can be zero only if $a_{31}=a_{32}=0$. Therefore, the discrepancies cannot be all zero unless the second set of primaries is identical with the original set and $r=R$, $g=G$, $b=B$ in general.

As a consequence of this demonstration, Wright's curves have been taken as the standard for the discussion because *by the method of determination* they automatically meet the requirements laid down. Dr. MacAdam's demonstration implies that they are the *only* ones that can.

III. EFFECT OF ADAPTATION LEVEL

Introduction

One of the most common visual experiences of everyday life consists of a change in the quality of colors with the amount of illumination. This phenomenon is particularly noticeable in work with color photography, in which the proper illumination may pass through a sharp optimum value. Since the practical aspects of the subject have been discussed very little in the literature of color vision, a qualitative review of the subject, combined with an attempt to link up the known facts with those of brightness and color constancy, may not be out of place.

Many people have observed that colors appear "more brilliant" in summer than they do in winter, especially in the northern latitudes. The same difference is observed between colors viewed on a clear and an overcast day. The effect appears to be due to the difference in the illumination levels. Perhaps the best known example and one which is familiar to all readers is the difference in the appearance of automobile colors (more particularly those used ten or fifteen years ago) when illuminated by sunlight and when seen on dull days. Many of these colors could not be distinguished from black even with a comparison black beside them unless the illumination was fairly high. In sunlight, however, the various colors were seen as well-defined hues, some of them having surprisingly high saturation.

Such phenomena give direct evidence that color as seen depends directly on the illumination intensity aside from all considerations of its spectral energy distribution. It is the purpose of this article to review several aspects of the dependence of color on illumination intensity and try to interpret the facts so that they will explain the observed results in color photography. It is intended merely to separate and indicate the variables in a qualitative way in the hope it may lead to fruitful quantitative work.

Photography, like painting, attempts to reproduce a given scene, not only in form and perspective but also in appearance. This means that, regardless of the nature, color, and amount of the illumination of the original scene (of which only part is recorded), the reproduction should result in the same visual perceptions in an observer, irrespective of the conditions under which it is viewed. Unlike painting, photography attempts to do this with a relatively inflexible technique in that individual areas cannot be arbitrarily modified. Obviously, a complete and rigorous solution of the problem presented is not to be expected, but it is, nevertheless, a practical problem which workers in color photography must face for a public unconcerned with the difficulties.

In the classical tone reproduction theory of black-and-white photography as developed by Jones and others,[1] this phase of photography has very properly been considered as external to the usually more important problem of obtaining a point-for-point match between the relative brightnesses of a scene and of its reproduction. Much of the literature on the practice of photography, however, deals with it under the heading "How to produce such and such an effect." A typical example is the use of infra-red photography to imitate night scenes. It is not usually realized that it is the dependence of the visual processes on the illumination level which is involved.

In color photography with its closer approach to the illusion of reality, the eye also demands a closer approximation to the true appearance of the scene photographed. In addition, the viewing of color prints under different illumination conditions produces greater apparent changes in the prints than is the case in black-and-white work.

While it is undoubtedly true that there is no general solution which meets all the requirements, it is equally true that the problem cannot be disregarded in any complete theory of color photography.

Nature of the Adaptation Process

The sensitivity of the eye to light is highest when it has rested for long periods in the dark. Sudden viewing of a light source under these conditions or sudden illumination of the room produces first a blinding sensation of glare with almost no vision of detail until the visual mechanism gradually adjusts itself to the situation. The time taken for this adjustment or adaptation depends on the average luminance of the scene. The final appearance of the scene to the observer, however, is more or less independent of this value over a wide range of intensities.

The total amount of light entering the eye is variously distributed over its sensitive surface. The greater part forms an image of the scene but considerable portions enter as stray light through the eyeball, are scattered by the media of the eye, etc., and represent non-image-forming light to which the eye also responds.

The eye itself is in almost continuous movement, seldom stopping for longer than a tenth of a second in any one position. Attempts to hold it motionless result in minute quivering motions. Perception takes place only after the eye has been relatively stationary for a brief interval, as can be seen from the absence of apparent detail in the stationary background when the eye follows a rapidly moving object. Continued fixation of the eye on a particular object, however, soon decreases the apparent contrast and, when vision is transferred to another point, produces "after-images" of the first object.

The various adjustments which take place tend to make the perceptions due to light of any luminances fall within a certain range of brightnesses. The range may be described as that from blinding brightness, through white and gray, to black (and correspondingly for chromatic sensations). These adjustments are known collectively as adaptation phenomena. They uniformly tend to produce the most favorable response condition of the eye for the viewing problem presented by the scene. An analogous process exists in photographic printing when the amount of exposure through a negative is adjusted so that the scene best fills the scale of the paper. As Schouten[8] has pointed out, the eye is used primarily as a "null-type" instrument to determine whether an object is lighter or darker than its surroundings, not as an absolute instrument to determine the illumination level.

As was seen in the previous parts of this article, three types of brightness adaptation must be distinguished. They are general, local, and lateral adaptation. All three must be considered capable of occurring locally in the eye in the sense that two areas of the same retina may differ in all three. There appears to be no adaptation of the eye *as a whole* except insofar as conditions external to the eye may produce such an effect.

From these considerations it is possible to piece together a picture of what is meant by the term "adaptation level," at least as far as the fovea is concerned.

The constantly moving eye itself adapts locally and rapidly, the extent depending on time of hesitation, intensities of the objects viewed, and the angles which the brighter parts of field make with the line of sight. This process and the general nature of ordinary scenes give rise to two general conditions which exist simultaneously.

From the results of a study of such scenes by Jones and Condit[21] it appears that the average outdoor scene has an average directional reflectance in the neighborhood of 20–25 percent and that the ratio of maximum to minimum luminance for scenes in daylight varies from 25 up to 750. The eye under these conditions is exposed for long times to luminances which fall within these limits. Unless it is fixed on some particular point of the scene it will be exposed successively to the various luminances and become adapted to some value lying between these limits. From such long time exposures recovery of sensibility is relatively slow. Therefore, since the whole eye is exposed, it will take up a sort of average adaptation condition corresponding to the term "general adaptation." Its actual value probably lies close to the average reflectance unless the eye for some reason remains for several minutes in a relatively fixed position.

In the process of looking at the scene, the eye views one object after another, stopping for brief intervals at each one. At each of these stops a readjustment of the adaptation takes place locally. From such adaptation recovery is very rapid, the exact rate depending on the time of exposure. If the interval is as long as a second or more, the observer may be conscious of "after-images" owing to the lag in recovery of the local sensitivities. This lag is particularly noticeable when there is a sudden increase in illumination level and before general adaptation has had time to become established. The various areas of the eye are hence in a ceaseless state of changing relative sensitivity.

In addition to the changes in sensitivity directly produced by the retinal image of the scene there is the third effect which plays an enormous part in the appearance of objects. If a moderately dark object is surrounded by considerably brighter ones, the sensitivity is not decreased at the light points in the image only, but this effect extends into the darker part decreasing its sensitivity also. This results in an apparent darkening of adjoining brightnesses and, in extreme cases, to a loss of detail in the darker objects. This "lateral" adaptation may be of large magnitude where considerable brightness differences are involved. Since it depends markedly on the relative areas and positions of the objects concerned it plays a large part in some of the failures of photographic reproduction.

All of these adaptation effects are steeply variable from one person to another. It appears, therefore, that the expression "adaptation level" represents an over-simplification of the situation. In ordinary usage it corresponds to general adaptation, but is frequently used, for example, in reference to the viewing of a small isolated patch of light, or to describe the state of an observer's eye at some time after conditions have changed. The expression, "as seen by the totally dark-adapted eye" is frequently encountered.

Any attempt to generalize the facts and deduce quantitative relations for the sensitivity characteristics of the fovea at any given time encounters numerous difficulties. At least three factors must be considered as dependent on the time. They are: the average illumination, the effect of small areas of high intensity, and the size of the object being viewed. They must be considered from the standpoint of least perceptible energy and least perceptible energy difference throughout the visible energy range. The ultimately desirable data would deal with visual contrast at all levels. Contrast, however, is a perception and as such may or may not correlate with definite luminance relations. For example, in the brightness constancy phenomena discussed in the first part of this article, what can be said concerning the contrast relations existing when a white object in shadow appears much brighter than a nearby black in sunlight in spite of the fact that the black has a higher luminance?

The usual technique for studying the effect of adaptation level is to surround a small field with a large, uniformly bright area and then to refer to the level in terms of the luminance of this surround. Except over very small fields, the sensitivity of the fovea is not made uniform by this technique, and for these fields lateral adaptation is at a maximum which varies with the degree of fixation of the eye. Schouten has shown, for instance, that for central fixation the sensitivity may be more than twice as great at the center of a six-degree field surrounded by a circle than it is near the boundary. It is doubtless a still higher ratio if points very close to the boundary

[21] L. A. Jones and H. R. Condit, J. Opt. Soc. Am. **31**, 651 (1941).

are considered, although constant viewing of the border would tend to reduce the foveal sensitivity to the proper level.

With these limitations on their extension to everyday vision, certain facts seem well established. Holladay and Stiles[22] have shown that the sensitivity of the fovea to luminance differences is markedly decreased by brighter surrounds, a slightly brighter surround producing nearly the full effect. In this work it was found possible to derive an expression from which the "equivalent background" effect of any visible bright area could be calculated.

Cobb and Moss,[23] among others, have shown that when the surround has the same luminance as the actual area numerous visual functions such as acuity, luminance sensibility, etc., are at a maximum. This accordingly usually represents the condition of maximum "seeing."

LeGrand[24] has shown that in the absence of any other illumination in the field of view a small light spot produces a light sensation at the fovea even when its image lies some distance from it. He finds that this sensation depends only on the intensity and the angular distance from the fovea. For very small angles the equivalent luminance at the fovea varies as the reciprocal of the third power of the angle, thus producing a sort of spreading of the image. For larger angles it varies with the inverse square and ultimately as the reciprocal.

LeGrand's work indicates an increase of illumination due to a light source off the line of sight. The work of Schouten, Holladay, and Stiles indicates a loss of sensitivity from the same cause. A photograph involving the perception of a black introduces both functions in an exceedingly complex way.

Abribat[25] tried to measure directly the just perceptible luminance in a somewhat complex field. He used a small field, graded in luminance in two directions, and surrounded by a uniformly bright field which completely filled the rest of the eye. He found that the minimum luminance perceptible under these conditions is a fairly simple function of the surround luminance. Above about 30 millilamberts, the ratio of just perceptible to surround luminance is approximately constant up through daylight luminance. The value of the ratio (computed from his published curve) is about 3×10^{-3}. Below 30 millilamberts, the ratio increases steadily, reaching a value of unity in the neighborhood of 10^{-3} millilambert, i.e., at this point the just perceptible luminance matches the surround. These figures are in general accord with the earlier, less complete findings of Lowry,[26] who used just perceptible binocular differences in a very small field with a very large surround.

These findings indicate that from daylight levels down to about 30 millilamberts the ratio between an intensity which has produced general adaptation and that which is just perceptible is about 300. Below this level, the ratio decreases continuously to the threshold of vision.

It remains to consider the intensity range upward from the adaptation intensity and to consider particularly the situation with respect to white surfaces in a complex field. In such a field the adaptation is almost certainly below that of white. White, however, as has been noted, is a perception which is not confined to surfaces whose reflectance is 100 percent and whose purity is zero, since very different surfaces may appear white under suitable conditions.

Little work seems to have been done in this field and at best only a guess can be hazarded as to the true situation. For surfaces perceived by reflection, brightnesses range through gray up to white and in the same field of view surfaces of many different luminances may appear white. One is tempted to suggest that the perception of white may be produced by any luminance appreciably above that producing the existing general or local adaptation. It is probable, however, that the adaptation level is usually much lower than this would imply and that white is more closely associated with the brightest nonselective diffuse reflecting surface in the immediate field of view.

Jones and Condit[21] found that the average reflectance of the scenes as measured was one-

[22] L. L. Holladay, J. Opt. Soc. Am. and Rev. Sci. Inst. **12**, 271 (1926); J. Opt. Soc. Am. **14**, 1 (1927); W. S. Stiles, Proc. Roy. Soc. **104B**, 322 (1929).

[23] P. W. Cobb and F. K. Moss, Trans. Illum. Eng. Soc. **23**, 1104 (1928); P. W. Cobb, *ibid.* **11**, 372 (1916).

[24] Y. LeGrand, Rev. d'Optique **16**, 241 (1937).

[25] M. Abribat, Réunions de l'Institut d'Optique **6**, 3 (No. 3, 1935).

[26] E. M. Lowry, J. Opt. Soc. Am. **18**, 29 (1929).

fourth to one-fifth the brightness of the average maximum insofar as the average has meaning in non-uniform illumination. Since it is this sort of integration of the light which determines the general adaptation it is of interest to see how far the maximum exceeded the average for some of the specific cases listed. The brightest ratio listed is more than ten times the average reflectance for that scene and the minimum is about one and one-half times. Since the brightest surface in each case could have been and probably was perceived as white, it is safe to say that whites can vary *at least* up to ten times the general adaptation. Incidentally, in uniform illumination, a reflectance of 10 percent is perceived as a very dark gray while $66\frac{2}{3}$ percent is very light, these corresponding to the adaptation levels with respect to the highest and lowest maximum brightnesses, respectively.

Taking Abribat's value for these levels as a ratio of 300 between adaptation level and the deepest possible black, the maximum ratio of white to average gives a possible brightness range of 3000 to 1. This implies that the latitude of the eye is this great without the help of local adaptation. It may, perhaps, be considered as a rough approximation of the possible range in daylight. At low levels it is much less than this, however. Assuming that the same physical conditions can hold at a low illumination level (for example, at one millilambert), then Abribat's results would suggest a maximum range of 1000 to 1 or less, and at 0.1 millilambert, the value would be of the order of 200 to 1. This is mere speculation, however. What Abribat's results do appear to show definitely is that, at a general adaptation level of one millilambert, luminances less than 1/100 of this will not be visible and that at 0.1 millilambert, brightnesses less than 1/20 will disappear. At low intensity levels, therefore, the range of visible details below the general adaptation level is greatly compressed causing a loss of details in the shadows. This must be reproduced photographically by an equivalent actual loss if anything like the same effect is to be reproduced by the print.

Nutting[27] has shown that the ratio of minimum perceptible radiation to just tolerable *glare* intensities is greater at low general adaptation levels. His ratios of adapting luminance to just perceptible, however, are of the same order of magnitude as those of Abribat.

These considerations have some interesting consequences with regard to the nature of perception of scenes.

The fact that a scene can be quite bright or quite dark with respect to the same adaptation level, depending on the relation of maximum and minimum luminance to the average, indicates that scene contrast and apparent scene brightness are not controlled entirely by the state of sensitivity of the eye. It suggests, in fact, that apparent contrast may be determined by the relation of maximum and minimum luminance to average luminance and that the brightness of a scene may be almost wholly controlled by the extent to which the brightest area is illuminated *above* the general adaptation level or by some similar function.

It may be suggested here that the concepts of light and dark as applied to a whole scene, such as a "dimly lighted interior," a "dark" day, a "bright" day, or a "brilliantly lighted" stage setting, are more directly associated with the *contrast* of the illumination of the scene than they are with the average luminance as such. This approach affords some promise at least of explaining the frequent total failures of photographic reproduction to recreate the effect of apparent lighting intensity. The brightness, of course, can be imitated to some extent in the printing process by making the print lighter or darker and so varying its level with respect to general adaptation. It cannot be made to appear to have *more* brilliant lighting (except as a logical fact rather than as a perception) by any technique which does not include *local* modification of the print.

So far as contrast is concerned, therefore, *eye sensitivity* data can yield only a factor for the eye which applies under certain conditions, and is roughly analogous to the γ of photographic sensitometry. If the condition is changed, however, the perceptions change according to some quite different function.

At present, this function appears to be unknown. Judging from experience with photographic printing in black and white, a factor would be introduced which would be quite high

[27] P. G. Nutting, Trans. Illum. Eng. Soc. 11, 1 (1916); *ibid.* 11, 939 (1916).

at the *lowest* intensities and quite low at the highest. A photographic print seen in the low levels of a photographic dark-room appears considerably more contrasty than at ordinary levels and is, in turn, more contrasty than it appears in full sunlight, especially if it is a low contrast print. If the print itself has an exceedingly high ratio of reflectances, the reverse may actually be the case. Contrast then appears to be about equal at all average luminances if good detail is just visible in the deepest blacks and both a good black and a good white are visible, in other words, if the "scale" of the eye is just filled. If detail is obviously lost in the shadows and they appear black, contrast will appear higher; if brighter whites or deeper blacks are obviously possible, contrast will tend to appear lower.

Wright's work, as well as Schouten's, shows that the various brightness adaptation phenomena of the fovea apply to colored as well as to neutral stimuli. Before discussing the relation of this phase of the problem to color photography, we shall consider the various ways in which colored stimuli are modified in their appearance by their physical intensities, both in absolute measure and in relation to their environment. Actually, the question is how is the perception of color by the eye modified by the stimulus itself and by its environment.

There are only three fundamentally different ways in which a colored area can occur in a scene. It may occur as a completely isolated patch, as in an optical instrument, with no stimulation of the eye in the region around the image. It may be surrounded by fields of any color under any color of illumination, and it may occur in a portion of a total field which is subject to a different illumination from the remainder of the field. Depending on the exact conditions, these three fundamental situations give rise to numerous differing series of color perceptions. The two cases with which we are chiefly concerned are as follows.

Case 1. Isolated Color Patches With No Surround Illumination

Variable Intensity

As the luminance of an isolated chromatic stimulus increases from very low values to very high values, the perceptions produced vary in all the attributes of color. Over a short region of intensities there is no hue. This is followed by a region of gradually increasing saturation and brightness in which hue is not, in general, constant. Above this range, hue and saturation change more slowly than brightness for a short time, followed by a region in which saturation decreases, again with considerable shift in hue as the intensity becomes more and more painful. At the lower end of the range in which color is seen, the brightness can properly be described as weak. Increases in intensity make the brightness higher and higher until it becomes unbearable. There is no luminance under these conditions in which the color appears dark, in the sense that a pure pigment becomes dark when it is mixed with black. Neither is there any region in which the color appears to change in the amount of white which it contains. At very high intensities there is a tendency for the hue to diminish and this diminution is described as a loss of saturation. The sensation, however, is that of being blinded for color rather than of admixture with white.

Case 2. Colored Area Surrounded by Daylight

The statements just given apply only to a single isolated stimulus under conditions (either monocular or binocular) in which there is no other stimulus present in the field of view. The presence of another stimulus of any area or nature changes this whole series more or less completely. Perhaps the simplest case, and the one having the greatest bearing on the present problems, consists of a two-part field in which a relatively small area subtending a few degrees at the eye is surrounded by a much larger field, both areas being uniform throughout, and the dividing line between the areas being sharp.

In such a two-part field, color perception takes on two new attributes which are usually described as "contrast effects." These attributes consist of apparent admixture of black with the less intense part of the field and of admixture of white with the more intense part. With the relative sizes of field mentioned, however, these effects are best seen when the central small area is made darker or brighter, respectively, than the surround and when the surround is made of white material. If the luminance of the surround

is fixed at a certain moderately high luminance and is of daylight quality, then the following sequence of perceptions takes place when the intensity of the central area is increased from zero. The area first appears black. This black is a darker and a more positive perception than absence of light without the surround. As the intensity is increased there is no change in the appearance of the black until a certain level is reached which is different for each surround luminance. At this level the black is slightly tinged with color. As the intensity increases further, this color increases and the percentage of black decreases until, at the same brightness as the surround, the black has disappeared completely and the stimulus is seen at its maximum color. Above this luminance, the effects depend markedly on the nature of the color, the size of field, etc. If the central field is fairly large, increase in its luminance above that of the surround appears to decrease the saturation of the color up to a certain point, above which saturation remains constant and further increases have the same result as though no surround were present. If the color sensation from the stimulus is saturated when the brightness matches the surround, it will retain most of this saturation when the level is too high to be affected by the surround. If the color is very desaturated, it may lose all trace of hue before it reaches this point. In this case, it does not look as though it were white or mixed with white but rather as a brilliant achromatic source of light. In general, hue varies noticeably from the point at which it appears admixed with black through to the highest luminances.

The limits of the range, from just perceptible hue in the black up to a brightness match with the surround, and from this point up to independence from the surround, vary tremendously with the luminance of the surround. For low luminance, both ranges are relatively short, the ratio of brightnesses being of the order of 10 or 20. For luminances of daylight magnitude, the lower range may be 1000 to 1 and the second shorter than this. At very high surround luminances, it may not be possible to tolerate the brightnesses necessary to make the central part independent of the surround.

It is important to consider more closely the changes in the appearances of the stimuli of various hues as they are made to pass through these and other series.

The simplest series from the standpoint of color attributes is the case of the isolated field which is varied in luminance only. Starting from zero with dark-adapted vision, the sensation is first achromatic, then gradually gains hue which then shifts with increasing luminance (the Bezold-Brucke effect). In general, throughout such a series the saturation changes very little except at the extremes of the range over which vision is possible. The series differs fundamentally from the others in the almost complete absence of the sensation of gray throughout the normal visual range. As the intensity changes, there is a change in the *brightness of the field only except* for the hue changes, which are usually hardly noticeable. The field at all times has the appearance of a chromatic color mixed with some white. Throughout most of the range, for example, an orange stimulus looks like some part of the spectrum in the region between yellow and red but varies in brightness. When an achromatic comparison field is introduced, the sensation series varies in a new attribute, that of admixture with gray. With some stimuli the presence of this new attribute produces sensations which are sufficiently different from any in the first series to have received special names in everyday speech. Perhaps the most distinctive of these is brown. Brown is the sensation produced by a stimulus corresponding to the orange-red end of the spectrum *when the eye is at a higher general or local adaptation level than that produced by the chromatic stimulus itself.* Hence it can be seen only when there is a *brighter* comparison field. The same thing holds for all colors containing gray. There are no dark colors in an isolated field, no "olive drabs," no khaki, no "navy blue" etc. All of these are a combination of the sensation produced by an isolated stimulus plus the sensation of gray, and the amount of gray depends not on the stimulus itself but on its brightness *relative* to that of another stimulus. In colorimetric terminology, the variable is relative luminosity. For chromatic stimuli, then, gray *is the perception* produced by relative luminance just as was found for neutral stimuli. The series of sensations generated by the stimulus field when the white surround intensity is varied is a

series in relative luminosity. At very low values relative to the surround, the perception is that of a black; at moderate values, color plus gray; and at a relative luminosity of one, no gray is seen. When the value exceeds unity, two possible types of change occur. Either the stimulus field looks like a light source, i.e., "glows," or it takes over the control of the adaptation of the eye and the "gray sensation" is transferred to the surround. In the first case, there is a transfer to the condition produced by an isolated field usually with a noticeable loss of saturation; in the second case, the perception is better described as an apparent admixture with white. Obviously, the two perceptions are simply different ways of looking at the same thing. The type of perception in a particular case is determined by the spatial and temporal characteristics of the fields as well as by the saturations and relative luminances.

The foregoing discussion involves a concept which has been mentioned previously but which is not customarily stated in these terms, namely, that of the *identity* of gray with relative luminance at a given adaptation.[28]

For the present study, this fact may well be worded as follows. Under conditions in which white paper or its equivalent is to be considered as white and the illumination is uniform, relative reflectance *with respect to this* white corresponds to a definite gray. Non-uniform illumination or the absence of high reflection surfaces will shift this value to an unknown extent but in a predictable direction.

Brightness Perception under Color Adaptation

In Part II of this article, careful consideration was given to the phenomena accompanying general adaptation of the eye to a colored stimulus. It was stated that brightness adaptation might be considered as a separate phenomenon and some evidence was given to support this contention. The relative brightness of chromatic stimuli was not considered although it was mentioned that the appearance of a blue surface was noticeably darker relative to red in artificial light than in daylight. This is a general phenomenon which may be summarized as follows. The relative brightness of colors in artificial illumination differs from that of daylight in the same way in which the light source differs from daylight. If the source is relatively weak in blue, the blues will be similarly darkened and similarly also, for any other region of the spectrum. So far as color adaptation is concerned, relative luminance may accordingly be determined directly by the usual I. C. I. calculations. The integral of the spectral energy distribution curve of the stimulus with respect to the standard luminosity function gives the relative luminance (the Y of the I. C. I. system) directly.[29]

A photographic print viewed by artificial light will then show the same loss of brightness in the blues as did the subject, regardless of the particular spectral absorption of the two blues, since the effect depends only on the source of light used. A color film which corrects artificial light so that the photography appears to have been done in daylight produces correct results. Again, however, it is not possible to reproduce the *effect* of artificial light unless the print *also* is viewed by artificial light.

Possible Calculations of Color Perception

The possibility of calculating the appearance of a given stimulus can now be considered in the light of the foregoing résumé. It is exceedingly important to the theory of color photography to be able to make such calculations, because it is only by the deduction of general laws which film may be made to obey that corrections can be introduced. Needless to say, such deductions have yet to be made in complete form. Their nature, however, can be suggested.

Any color stimulus may be specified by three variables, such as those of the I. C. I. system. The general or local color adaptation may also be specified by three. The general brightness adaptation level requires one (the average luminance of the scene at the moment). The spatial distribution of the stimulus with respect to its surround requires three relative tristimulus values to introduce the effect of lateral adaptation. To these it

[28] D. B. Judd, Am. J. Psychol. **54**, 289–294 (1941).

[29] In a private communication, Dr. D. B. Judd points out that for intense adapting stimuli there is evidence that changes are produced in the luminosity function. See also N. T. Federow and V. J. Federowa, Zeits. f. Physik **57**, 855 (1929); *ibid.* **62**, 834 (1930).

may be necessary to add the luminance which the observer considers to be white.

From these ten variables the appearance of a stimulus should be calculable in terms of a known and standard group of stimuli (an area of definite size with a given surround at a given general intensity level, for example), provided the functions connecting the variables are known. Judd[2] has attempted an empirical solution of this problem and has deduced an equation by which the hue, saturation, and lightness of a given colored paper may be calculated under certain circumstances.

Fortunately for color photography, however, the problem presented is simpler than this. If a color process can be made to give a point-for-point match of a given scene considered as a group of physical stimuli with respect to the eye, then it is the *relative* effect of the remaining seven pairs of variables in viewing the scene and the reproduction which must be considered. Because of the phenomenon of color constancy, a good approximation for most cases can be obtained by eliminating the three variables of relative color adaptation, although the illumination must be considered from the standpoint of the relative luminosity of the colors. The phenomenon of lateral adaptation is then governed by the relative sizes (visual angles) of the scene and the reproduction. If sufficient data on this function were available, it might be possible to change a process to allow for it. The effective adaptation level of the eye, however, depends on the scene as a whole and on the way the observer looks at the scene. In like fashion, it depends on where and how he looks at the print. Only by assuming that the observer will view the reproduction under conditions which fall within certain permissible ranges can any worthwhile modifications be introduced. The requirement that the luminance being taken as white must be specified is easily met but places a definite limitation on possible processes. Objects which are to appear white must be white, i.e., must match a piece of paper which is obviously white. In the case of a paper print, such a comparison can be supplied in the form of a white border.

No attempt will be made to deduce any of these relations, if indeed, they can be deduced from existing data. There is no fundamental reason, however, why they cannot all be determined experimentally and the consequences for reproduction studied with respect to possible inclusion in the theory of color photography.

Experience over some years in the mass production of color photographs both for general amateur use and for the more exacting demands of the professional field has indicated the relative importance of many of these variables. Many effects which force themselves on the attention are based on visual processes which have been very little discussed in the literature of vision but which have been considered here at some length. The remainder of the article will discuss the observed photographic effects, pointing out either their connection with the foregoing discussion or stating the problems involved in such a correlation.

Photography at Different Luminance Levels

In black-and-white photography in which brightness is the chief variable, the contrast of a picture may be modified by changing the contrast factor or gamma of the process. This can be done in several ways, such as increasing the time of development of the negative, choosing a more contrasty printing paper, etc. Since relative brightness is the variable chiefly affected, this procedure is entirely satisfactory. In color photography, however, a change in the contrast factor or gamma of the process modifies not only the brightness relations but also the saturation of the colors. It will be noted from the preceding discussion that the contrast of a scene as a visual phenomenon varies independently of saturation. Hence, any particular color process can be used only at one gamma, the actual value being determined primarily by the level of saturation reproduction required by the particular process. This fact, coupled with the fact that subtractive processes in general have a higher gamma for their reproduction of neutral scales than for colors, has important consequences. The first of these, of course, is that all subjects must be handled by means of a single contrast factor, i.e., by a fixed photographic color process with exposure as the only variable.

The range of exposures to simple stimuli which can be reproduced by a given photographic

material is limited, perhaps of the order of 300 to 1. By means of the time factor, however, and because of the fact that in photography the product of intensity and time may be made roughly constant, any given brightness in a scene may be reproduced at any desired density within this range. It is seen at once that the *time of exposure* in photography plays a role which, with respect to the photographic material, is very similar to that played by adaptation processes in the eye with respect to the range of customary perceptions. Both processes place the effect of the stimulus within a certain range. There are also other similarities. Underexposure of photographic materials causes loss of details in the shadows in much the same way that the eye loses them through loss of low brightness sensitivity at low levels. Overexposure causes a loss of detail in the highlights similar to the effect of glare in the eye. Here, however, the similarity ceases.

It has just been shown that in vision any particular scene ranges from a brightness which will be seen as white to one which will be seen as black. Although, in any *one* scene these have definite relations to the general state of sensitivity of the eye at the time, this relation is not general, that is, it is different for every scene. It is also probable that it is just these relations that determine the general *impression of brightness* experienced by the observer. Since this impression, among other things, must be reproduced by the photograph the problem may be stated as follows: How can any scene be so photographed that the brightnesses which were perceived as white and black by the observer under one set of conditions will again be so perceived under *any* viewing conditions?

Since reflectances of the order of 80–100 percent and approximately zero colorimetric purity are seen as white under nearly all conditions, the first requirement which must be met is, as noted before, that whites in the scene must be reproduced as white areas on the print. If it can be assumed that all prints are to be viewed under *normal* illumination levels, i.e., not lower than 15–20 footcandles, then the density which will be seen as black may be approximately specified. *All scenes, then, must be fitted to these values with respect to the observer at the time the picture was taken.* In black-and-white work, this can be and is done by modification of the contrast factor. In color work, this variable cannot be used in the same way. Before discussing what can be done, however, we shall consider further the relation of maximum and minimum brightness to the average.

If the impression of general brightness is produced by the relation between the luminance of white and the average luminance, this relation should be retained in the reproduction. However, if the luminance range of all scenes is brought to a constant value, this condition can only be met by a *distortion* of the relative brightness reproduction scale, unless the observer views the print under conditions such that the *average for his surroundings* has the same relation to white as it did when he viewed the original scene. There are two possible solutions of this difficulty. Either the print and its mount may be made to control the general adaptation level of the observer, or distortions can be introduced based on some assumed average viewing condition. A point-for-point relative luminance reproduction must be seen under the same *average brightness* conditions as the original subject to give the same impression. Conditions under which the reproduction does control the observer's general eye sensitivity will be discussed under "Effect of Viewing Conditions"; the introduction of equivalent distortion for an assumed viewing condition will not be discussed here.

Returning to the problem of fitting any scene into fairly definite density limits in color photography, it should be noted that this must be done without modifying the contrast factor of the color process *as such*. Although the saturation of the colors is controlled by the contrast of the process, this is no longer true after the reproduction has been obtained. By superimposing a neutral image (black-and-white) of the desired contrast (positive or negative), the brightness reproduction may be modified to any desired extent without changing the chromaticity of any of the colors. Such superimposed images are known as "tone correction *masks*" and are of wide possible application. Since the relationship of density to exposure may be varied at will, it is possible by this technique to control the "curve shape" of the color process completely without affecting its color reproduction. The use

of these masks, however, is limited by the fact that they have been successful only over color transparencies and, of course, are of use only during a copying process. They do not, in general, improve the appearance of an original but only of the print made through the original and the mask. Numerous practical difficulties are also encountered. With these masks, it is possible, however, to meet the requirement of printing any transparency so that any two densities in the transparency correspond to black and to white in the print and at the same time to make the average reflectance fall at almost any desired point without making any change in the final printing process. Their use is almost essential for some types of work but the control which can be utilized is limited by economic considerations and by the fact that a more serious difficulty demands a somewhat different type of mask.

Even after attaining the best possible brightness masking, the white of the print is still no brighter than a white in the surroundings. Special viewing conditions, then, are necessary in order to give the desired effect in this case also.

In order to reproduce the effect of general adaptation level, therefore, it is necessary first to expose the picture so that shadow detail is represented only to the extent that shadow detail was visible at the time of viewing and then to print the transparency so that one end of the brightness scale is black and the other end of the scale is a true white.

This statement, however, introduces another complication because the eye is capable of rapid local adaptations. The effect of these adaptations is seen most clearly in large shadowed areas. There are three ways in which such areas may be perceived. They may be perceived (1) as part of the scene as a whole; (2) with the intent of seeing as much of the detail as possible within the shadow; or (3) overlooked completely as in the perception of the continuity of a single surface in shadow. In the first case, the eye does not rest on the shadowed area, the perception of it as an area is largely peripheral, and the sensitivity utilized is that of the general adaptation level. Careful study of the shadowed area involves long direct viewing with consequent relatively large foveal local adaptation to the level of the shadow and hence a greater perception of detail. The mechanism for the effects of brightness constancy referred to in the third case is not clear. The shadows appear much brighter than their relative luminance would lead one to expect.

In order for a reproduction to produce exactly the desired local adaptations in the observer it is probable that not only would it have to appear under identical visual angles with the subject but might also have to match the subject for luminance. The usual solution of this problem, however, is either to illuminate the shadows in the original subject so that after loss of effect the reproduction looks like the subject, or (what amounts to the same thing) to control the exposure locally in the printing operation ("dodging") to produce the required density difference.

Relative Brightness Reproduction Errors

In addition to the relative brightness errors which have been mentioned and which are due to the intensity and geometrical distribution of the illuminant, there are several others which are due to the defects of the color processes themselves or to the conditions under which they are used.

The first of these, as already noted, is the failure of any process to record the differences in relative brightness of colors when seen under chromatically different illuminations. The defect so far as brightness is concerned may be summarized as follows: When a color process is adjusted to correct for an illuminant, colors in which the source is relatively deficient will reproduce as too bright compared to their appearance to an adapted observer. For any single case, brightness corrections can be made which depend directly on the color by using colored light in making the brightness masks mentioned earlier. They then become what might be called "color brightness" masks.

More important for practical purposes because of their magnitude are the errors of all subtractive processes which are due to the general absorption of the eyes used. If all dyes in a color process contained the same amount of black, there would be a higher contrast factor for neutral stimuli than for colors, and neutral

areas would reproduce darker in color than in the subject. In the usual process there is more "impurity" in the cyan dye than in the magenta and least of all in the yellow. The result is that blue, cyan, and green colors tend to be much darker than the reds, oranges, and yellows. Under certain conditions this effect can be almost eliminated by the use of a "brightness" mask exposed to a definite color of light and developed to a definite gamma. This is known as a "color-correction" mask. Under certain favorable conditions, it may be combined with relative color brightness and contrast masks but the limitations are rather severe. In practice, this mask with or without modification to improve highlight contrast is the only one used.

The theory of masking for the various brightness and color defects of processes is exceedingly complex and difficult to explain. Only part of the theory has been published (Miller[30]), and space does not permit further consideration here. Experience has shown that the relative brightness errors are among the most serious produced by subtractive processes. Masks, or techniques which give the equivalent of masking, are almost necessary for their operation.

Effect of Viewing Conditions

It is interesting to find that many of the general defects of color processes may be overcome by the proper choice of viewing conditions. Although it is hardly possible to take advantage of this fact with a process designed for general use, anyone desiring to display a reproduction to the best advantage should take such factors into careful account.

The most favorable condition under which any type of picture is seen is exemplified by projection of transparencies in an otherwise completely dark room. Under these conditions, the screen acts as the equivalent of an isolated patch of color with a completely dark surround. The eye accordingly adapts locally, and to some extent generally, to the brightness *and average chromaticity* of the image. This adaptation produces two results, both of which are desirable: (1) A maximum amount of *gray* is *removed* from all the colors and (2) any slight lack of color "balance" in the process is corrected by the eye. Both of these effects make the reproduction appear *more correct* than it really is.

Under dark-room projection conditions, the over-all density of the transparency makes little difference, so far as faithfulness of reproduction is concerned. Most projection, however, is done by lenses which do not give a black of as low luminance as the rest of the room. That is, the darkest part of the projected picture is not as dark as the luminance perceived as black by the observer. Under these conditions, improvement in the contrast of the pictures can be obtained by decreasing the projection intensity or by surrounding the picture with a border of uniform brightness.

The effect of a surround on a projected picture is to decrease the sensitivity of the eye by affecting both the lateral and local adaptation processes. This decrease in sensitivity raises the intensity which is required to appear just brighter than black and gives a much more convincing reproduction of dark scenes which, without surrounds, permit such a great rise in sensitivity that they do not appear to contain black areas at all. The surround also acts in interesting ways as a fixed reference point both for chromaticity and for brightness. In theory, both of these effects are desirable, since they tend to extend the possible range of average brightnesses which may be projected. However, the presence of a fixed reference point also eliminates all the visual correction for light source and process balance errors. The use of a low intensity surround, therefore, while ultimately desirable, does not always give better results in projection.

Transparencies viewed over an illuminator in a partially darkened room are similar in appearance to projected images except for the generally greater luminance range due to absence of projection lens flare. The beauty of transparencies viewed in this way is usually ascribed to the greater density range available in transparency materials as compared with that of reflection prints. It is apparent that this is only part of the story. If an illuminator is so adjusted that the luminance of a white in the transparency exactly matches that of a sheet of white paper lying in room illumination, the transparency will usually appear less satisfactory than a reflection

[30] C. W. Miller, *Principles of Photographic Reproduction* (The Macmillan Company, New York, 1943).

print. Transparencies appear better primarily because they are customarily seen in such a way that their whites are very bright compared to the whites in the surrounding objects, and hence have a high white brightness relative to general adaptation. This permits the reproduction of scenes of much greater average brightness, removes the greatest percentage of gray from the colors, and, in general, greatly extends the usable range of the process. In these respects, however, it is not superior to projection in a darkened room except insofar as the greater effective luminance range plays a part. The color temperature of an illuminator is more noticeable as such, and transparencies so viewed must be more nearly correctly balanced.

A transparency viewed over an illuminator in a well-lighted room represents a transition case between dark-room projection and the viewing of reflection prints. Provided the highlights of the transparency on the illuminator are brighter than a white in the same position with respect to room illumination, there will be enhancement of whites in the reproduction and a longer scale of visible detail. The general adaptation level, both for color and for brightness, however, is now set by the room-lighting conditions rather than by the picture or by the illuminator (assuming that the whole surface is covered). The color of the illuminator must now approach rather closely therefore to that of the room illumination and the transparency must not be off-color by any great amount. The fact that the white is represented by a luminance which is usually considerably higher than is possible in a reflection print means that the picture as a whole has lost a considerable amount of the added gray owing to the general absorption of its dyes.

There is, finally, the most important case of straightforward reflection color prints viewed in normal surroundings but under all the very great variations in luminance and chromaticity which the word "normal" implies. Provided the illumination is uniformly distributed, all the visual effects operate in the direction of recognizing with maximum precision any faults which the print may have. Experiments indicate that under such conditions differences in hue which are just perceptible in the best two-part fields are easily perceived.

In a reflection print to be viewed "in the hand," then, everything in the process must be as nearly correct as possible. Whites in the original must be reproduced as white in the print. The average reflectance must be appropriate to such comparison surfaces as the observer's hand, etc., and color balance must be exactly correct or at least consistent with the subject matter of the picture. Needless to say, these conditions are not all met by any existing color process. The chief failure is that prints are too dark when hue and saturation are satisfactorily presented. This fact, coupled with the difficulty of maintaining exact control of color processes, is at present more important in making reflection prints than any of the other variables discussed. There are several ways of viewing prints, however, which markedly improve their appearance.

A dark print, by definition, is one which contains too much black in the reproduction of each color. The effect of this black content, however, is to decrease the "brilliance" of the print. In particular, such a print lacks what is known in the trade as "carrying power." This interesting term is difficult to define accurately and yet expresses a valuable property of a reproduction. If an observer stands close to a large print, the general and local adaptation of his eyes is controlled by the reflectance of the print. The eye then *tends* to see the print with as little gray content as it would in dark-room projection and whatever is represented as white *tends* to appear satisfactory. As the print is seen at greater and greater distances, however, the print has less and less effect on adaptation and the brightnesses are seen more with respect to their general surroundings. The print will "go dark" if the average of the surroundings is appreciably brighter than the print. Such a print is said to have "poor carrying power" because its effect does not carry to as great a distance from the print. For obvious reasons, the term applies equally well to the illumination changes which the print will tolerate. A dark print with poor whites decreases in brightness *faster* than its surroundings as the illumination decreases, at least at low illumination levels. A print with low carrying power, therefore, does not show up as well in a poorly-lighted part of a room. Carrying power of a given process could probably be expressed in the form

of a function containing the reflectance of maximum white, average reflectance, and saturation for the best print that could be obtained by that process. It is one of the few bases on which two fundamentally different color processes could be compared.

From a fixed viewing distance, the size of a print may have several effects. In general, the larger the area of a given color the more brilliant this color appears to the eye. In general, therefore, the larger the print the more brilliant will be the appearance of its colors. This is due partly to the increased control of adaptation given by the larger visual angle mentioned above, and partly to the *decrease* in the simultaneous contrast effects in the large areas from the usually less saturated surrounding color.

It is the usual aim in making a color print to approximate as closely as possible to the relative saturations and relative brightnesses of the colors of the original, provided other conditions have been met. In the original scene, the relative brightnesses were controlled very largely by what has been called "lateral adaptation" effects. These same effects take place in the print, their magnitude being controlled by the actual relative luminances and the visual angles subtended. Most processes give lower saturation and lower luminance on a point-for-point basis. The print then should be seen at a viewing angle which is at least as large as that subtended by the subject and larger angles may be necessary.

Some improvement can be made in the brightness of the print itself by placing it on low reflectance surfaces such as dark gray mounts and the like, the effect being similar to a local change of illumination. Great changes in the amount of gray in the colors may be produced by such mounts. It is thus even possible to obtain a fairly convincing reproduction of scenes containing brightnesses above those of the viewing conditions (sunlight on the snow, etc.). In spite of this fact, however, most people prefer white borders around prints rather than black. Presumably, association and similar phenomena are involved since a person stating preference for a white border will usually admit that the color reproduction is better on the black.

It is thus seen that color photography treads a precarious pathway through the visual processes. It is hoped that sufficient evidence has been presented to show that more is involved than the spectral distribution of the color and the relative sensitivities of the emulsions employed in its reproduction by a given dye system. Little has been added to the existing theory of color photography except to point out that much must be added before it will result in a workable scheme of operation. It is hoped that the visual phenomena to which color photography has thus called attention will receive more of the consideration which they deserve.

In closing, the author wishes to express his great debt to the workers in many fields who have made this article possible. Particular acknowledgment is made of the valuable comments and criticisms of the manuscript made by Dr. D. B. Judd of the National Bureau of Standards and by Dr. D. L. MacAdam of these Laboratories. Many others have discussed parts of the paper as well as the lecture and demonstrations given before the Optical Society on March 5, 1943, and considerable modification of the text has been introduced because of these talks.

New Institute for Fundamental Research in Glass

REPRESENTATIVES of leading companies interested in the development and use of glass met at the Hotel Biltmore in New York on August 17 under the leadership of Dr. W. A. Weyl, Professor of Glass Technology at Pennsylvania State College, and selected an organizing committee to prepare a charter for a new Institute of Glass Technology.

According to Dr. Weyl, the original sponsors of this institute are the American Optical Company, the Bausch & Lomb Optical Company, Corning Glass Works, Eastman Kodak Company, and the Pittsburgh Plate Glass Company. It is expected that membership in the new Institute will be open to all concerns interested in glass research.

Plans for the Institute include the establishment of laboratories whose location will be decided upon at a later meeting of the organizing committee.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 33, NUMBER 11 NOVEMBER, 1943

The Herschel Effect in Relation to a Critical Size of Development Centers*

G. KORNFELD AND T. H. JAMES
Kodak Research Laboratories, Eastman Kodak Company, Rochester, New York

(Received September 10, 1943)

The Herschel effect, on investigation at different density steps and with varying times of irradiation, proves to be a reaction which starts in each development center upon the absorption of one quantum. At low initial densities, a large number of the developable grains revert to the undevelopable state upon the removal of one or two silver atoms per center. Results obtained at different times of development in a standard hydroquinone-elon developer point to the conclusion that the "critical size" of the development centers remains essentially unchanged during the first few minutes of development, while at longer times of development centers of smaller size contribute to the developed density.

INTRODUCTION

VARIOUS attempts to determine the "size" of a latent-image nucleus which is just sufficient to initiate development have led to divergent results.[1] It is conceivable that the addition of *one* silver atom to a given nucleus will suffice to transform the grain which contains that nucleus from the non-developable to the developable state. It is also conceivable that no sharply defined "critical" size exists for a given nucleus, but that the addition of one or a few silver atoms to a nucleus will merely increase the rate at which development of the grain begins. This question can be examined experimentally if, by some suitable means, the size of the development centers produced by a given exposure, or the size of a statistically controllable fraction of them, can be reduced by *one* silver atom. The effect of the time of development upon the number of developable grains should then throw further light upon the relation between critical size and development time.

An experimental method of producing the desired decrease in the size of the development centers is available by using the Herschel effect, wherein a decrease of developable density is produced by exposure to red or infra-red radiation. In the Herschel effect, only radiation which is absorbed by the latent-image centers themselves is active. The absorption of one quantum by a silver nucleus primarily affects only one electron. If the electron leaves a center at the surface of a grain, a silver atom is converted into a silver ion, which then moves away from the center.[2] There is no reason to suppose that the neutral nucleus remaining should become unstable merely because it contains one silver atom less.

In the following section, some results will be given of an application of this method.

THE HERSCHEL EFFECT AS A "FIRST-ORDER" REACTION

An examination of published data on the Herschel effect[2–4] shows that, for the lower primary exposures to white light, the Herschel effect starts without an induction period. This absence of an induction period is quite significant for the purpose of the present investigation, as it indicates that at least some of the initially developable grains have been converted into non-developable grains by the action of one quantum per grain.

In the simplest statistical treatment of the Herschel effect, it may be assumed that N developable grains each contain one development center, and that the various development centers have roughly the same projective area. (The effect of grains which contain more than one center each will be treated subsequently.) If the N centers have absorbed s quanta, k centers will each have absorbed *r quanta or more*, and the fraction k/N is given approximately by the

* Communication No. 879 from the Kodak Research Laboratories.

[1] Cf. T. H. James and G. Kornfeld, Chem. Rev. **30**, 1 (1942).

[2] J. H. Webb and C. H. Evans, J. Opt. Soc. Am. **28**, 249 (1938).

[3] W. Leszynski, Zeits. f. wiss. Phot. **24**, 275 (1926).

[4] A. P. H. Trivelli, J. Frank. Inst. **207**, 765 (1929).

statistical equation:

$$k/N=1-e^{-s/N}\sum_{i=0}^{r-1}\frac{1}{i!}(s/N)^i. \qquad (1)$$

Now, if r is the number of quanta necessary for the reversal of a center, k/N will then be the fraction of the initially developable grains which have reverted to the undevelopable state. From Eq. (1) it follows that, when r is taken equal to one, the fraction k/N is given by the equation:

$$k/N=1-e^{-s/N}. \qquad (2)$$

Equation (2) is, in form, a simple first-order equation which possesses no induction period. If, however, r is taken as two or greater, Eq. (1) shows a definite induction period which increases with r.

In order to see how far Eq. (2) holds in actual practice, some Herschel effect measurements were made upon a contrast motion-picture positive film of nearly uniform grain size. The uniform grain size allows the substitution of optical density for actual grain count in the experimental procedure, since the two are approximately proportional for moderate densities. A strip of film was exposed to white light under standard conditions (1 second in an Eastman Type IB sensitometer, with a 500-watt lamp corresponding to 3000° K; step tablet with log E intervals of 0.15; log E at the eleventh step is 0.66 meter-candle-second). One-half of the strip was then exposed to infra-red radiation and the other half shielded for comparison. For the infra-red exposure, a Model X sensitometer was used, with the lamp current kept constant at 10 amperes. To insure the absence of any visible light, two Wratten No. 87 filters and one No. 25 filter were inserted in the light path.[2] The exposure times varied from 10 minutes to 18 hours. Development was carried out in a hydroquinone-elon developing bath (D-16) for 4 minutes at temperatures between 17.2° and 17.8°C. The strips were finished in the normal manner and densities read on a Capstaff densitometer employing a photronic cell.

TABLE I. Dependence of the Herschel effect upon the time of infra-red exposure.

t	$\log a/(a-x)$	b
10	0.02	0.002
20	.04	.002
30	.06	.002
40	.08	.002
60	.096	.0016
135	.164	.0012
240	.294	.0012
1080	.520	.0005

Equation (2) may be converted into a more directly applicable form by replacing k and N with the corresponding densities and substituting for s/N the exposure to infra-red light in the form bt, where t is the duration in minutes. In logarithmic form, Eq. (2) then becomes

$$\log a/(a-x)=bt, \qquad (3)$$

where a is the initial density, x is the *decrease* in density produced by the infra-red exposure, and b is a constant.

After 10 minutes of exposure to infra-red, the densities had changed only in the second decimal, so that no great accuracy could be expected, yet in one series the values for $\log a/(a-x)$ were found to be fairly constant: 0.03, 0.02, and 0.02, corresponding to three values for a ranging from 0.30 to 1.12. After 60 minutes' exposure, the values for $\log a/(a-x)$ for the values of a in the range 0.32 to 1.30 were 0.09, 0.09, and 0.09, respectively; after 135 minutes' exposure in approximately the same range of densities, the values were 0.17, 0.17, and 0.12, respectively.

These results justify the assumption that the density decreases proportionally to the decrease in the number of effective centers, and the dependence upon the time of infra-red exposure can be evaluated. In Table I the exposure times in minutes are given in the first column and, for the density range previously considered, the corresponding *average* values for $\log a/(a-x)$ are given in the second column. The third column gives the values for b according to Eq. (3).

The fact that b remains essentially constant until the density has dropped to about four-fifths of its original value shows that a considerable portion of the grains originally developable have reverted to the undevelopable condition by the absorption of only *one quantum* per development center. If all the grains required but one quantum each, the formula should hold throughout the entire range of exposures, but this is not the case. Some grains evidently contain centers which can be destroyed only by absorption of two or more quanta. Some grains quite probably contain two

TABLE II. Developed densities, with and without the Herschel effect, for various times of development.

2 min.			2½ min.			3 min.			3 min.		
a	$a-x$	$\log\frac{a}{a-x}$	a	$a-x$	$\log\frac{a}{a-x}$	a	$a-x$	$\log\frac{a}{a-x}$	a	$a-x$	$\log\frac{a}{a-x}$
0.05	0.02	0.40	0.04	—	—	0.05	0.02	0.40	0.05	0.01	0.70
0.11	0.05	0.39	0.11	0.01	1.04	0.16	0.03	0.73	0.15	0.03	0.70
0.31	0.09	0.54	0.33	0.09	0.56	0.40	0.11	0.56	0.38	0.11	0.54
0.65	0.22	0.47	0.71	0.22	0.51	0.75	0.28	0.43	0.68	0.23	0.47
1.10	0.49	0.35	1.18	0.50	0.37	1.26	0.56	0.35	1.22	0.54	0.35
1.50	0.95	0.20	1.74	0.99	0.25	1.78	1.07	0.22	1.73	1.07	0.21
1.89	1.38	0.14	2.30	1.54	0.17	2.27	1.63	0.14	2.26	1.57	0.16
2.22	1.89	0.07	2.74	2.16	0.11	2.69	2.19	0.09	2.68	2.14	0.11

4 min.			6 min.			10 min.			15 min.		
a	$a-x$	$\log\frac{a}{a-x}$	a	$a-x$	$\log\frac{a}{a-x}$	a	$a-x$	$\log\frac{a}{a-x}$	a	$a-x$	$\log\frac{a}{a-x}$
0.08	0.02	0.60	0.13	0.06	0.34	0.18	0.09	0.30	0.26	0.13	0.30
0.23	0.08	0.46	0.30	0.12	0.40	0.45	0.18	0.40	0.59	0.30	0.29
0.52	0.18	0.46	0.64	0.22	0.46	0.81	0.43	0.28	1.09	0.57	0.28
1.01	0.40	0.40	1.14	0.48	0.38	1.41	0.68	0.32	1.77	1.08	0.21
1.55	0.75	0.32	1.78	0.95	0.29	2.12	1.25	0.23	2.44	1.67	0.16
2.18	1.28	0.23	2.38	1.62	0.17	2.79	1.97	0.15	—	2.41	—
2.80	1.94	0.16	2.92	2.24	0.12	—	2.61	—	—	—	—

or more development centers initially. These grains, which will require at least "two or more" quanta in order to revert to the non-developable state, likewise will not follow the one-quantum formula. The fact that this formula does not hold for the longer infra-red exposure times thus confirms expectations from primary exposure theory.

CRITICAL SIZE OF THE CENTERS AT VARIOUS DEVELOPMENT TIMES

The extent of the Herschel effect was determined for various times of development. The experimental procedure was the same as that previously described. Four hours was chosen as a suitable time for the infra-red exposure, since it was desirable to produce a density decrease which was large compared with the margin of experimental error. Table I shows that an exposure of this duration does not result in too great a deviation from the one-quantum formula. The temperature of the developing bath was kept at 20°C and the times of development were varied from 2 to 15 minutes. The fog densities were negligible for short times and reached only 0.1 at 15 minutes. Consequently, the densities could be corrected by simple substraction.

Table II gives the corrected densities of eight series of experiments. Densities corresponding to the same intensity of standard white-light exposure are arranged in one horizontal line. The two series at 3 minutes' development time are already satisfactorily reproducible. Values for the lower development times are not so reliable.

The curves for the total densities at various development times are shown in Fig. 1. Up to 6 minutes' development time, the straight-line parts of the curves converge at a point on the log E axis. The increase in density with develop-

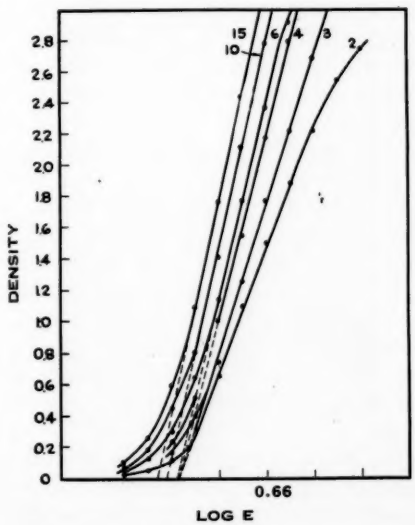


FIG. 1.

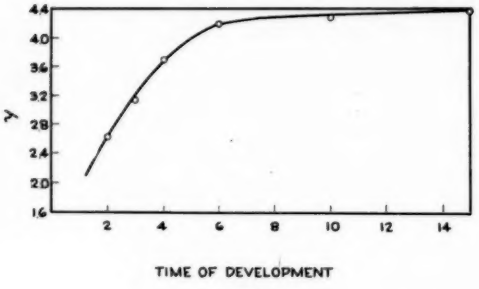


FIG. 2.

ment time is proportional to the density. The curves at 10 and 15 minutes' development time, however, do not converge at the same point, thus indicating a time region where the increase in density is less than proportional to the density. Gamma is approaching a maximum in this region, as shown by the gamma-time of development curve in Fig. 2.

The values for $\log a/(a-x)$, as shown in Table II, are fairly constant for higher densities up to 6 minutes' development time; at 10 minutes they show a decrease, which is even more pronounced at 15 minutes' development. The decrease at lower densities is partly due to changes of the photometric equivalent and to the necessarily large margin of error at very low densities. This is shown in Table III, where a, $a-x$, and $\log a/(a-x)$ have been determined in terms of milligrams of silver per 0.826 sq. cm for 3 minutes' and 10 minutes' development. The values are corrected for fog silver which amounts to 0.0046 mg for 3 minutes and 0.0108 mg for 10 minutes.

The values of $\log a/(a-x)$, which are equivalent to values of b when t is held constant, give a measure of the fraction of the grains which obey the one-quantum rule. The preceding tables show a general decrease of this value with increasing primary exposure above the thirteenth step. This result would be expected, since with increasing primary exposure the fraction of grains which have centers beyond the critical size or which have more than one center per grain will increase.

In the range of lower exposures, the variation of $\log a/(a-x)$ with development time may be regarded as significant if it is assumed that the infra-red exposure does not affect the nuclei in any other way than by decreasing their size. If the critical size of the centers was completely independent of the development time, this value should show no variation. Table III shows a small decrease in this value as the development time is increased from 3 to 10 minutes. Hence, development to 10 minutes consists not only of the completion (or near completion) of development in the grains which were already started at 3 minutes, but also of the commencement of development in new grains. The values for $\log a/(a-x)$ after 6 minutes' development, however, (Table II) are still in fair agreement with the values at 3 minutes. The development of grains already started is therefore the predominant factor in the early stages of development. The relative proportion of the two phases of development, however, should vary considerably with the composition of the developer.

A "critical size" of the development center which is so well defined that a decrease of one silver atom can seriously impair the ability of the center to initiate development does not seem compatible with the assumption of a large size. Rather, it suggests centers containing only a small number of silver atoms. It should be emphasized, however, that the exact critical size may vary from center to center, and probably depends upon the structure of the centers in question.[5]

TABLE III. Determination of developed silver with and without the Herschel effect, at two different times of development.

3 min.			10 min.		
a	$a-x$	$\log \frac{a}{a-x}$	a	$a-x$	$\log \frac{a}{a-x}$
0.0100	0.0040	0.40	0.0205	0.0100	0.31
.0221	.0076	.46	.0558	.0224	.40
.0462	.0162	.46	.1088	.0470	.37
.0859	.0353	.39	.1714	.0856	.30
.1339	.0678	.30	.2382	.1447	.23
.1811	.1110	.21	.3021	.2152	.16
.2126	.1736	.09	.3465	.2789	.11
.2664	.1989	.13	.3815	.3341	.06
.2878	.2585	.05	——	.3604	—

ACKNOWLEDGMENT

We are indebted to Mr. A. Ballard and Mr. B. Hutchins of the Chemistry Department for carrying out the silver determinations.

[5] S. E. Sheppard, Phot. J. **71**, 313 (1931).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 33, NUMBER 11 NOVEMBER, 1943

The Effect of Hue on Dark Adaptation*

EARL M. LOWRY
Eastman Kodak Company, Rochester, New York
(Received September 7, 1943)

The writer believes that the reported reduction in time necessary to dark-adapt the eyes by wearing red goggles for a period of approximately 30 minutes is due to the much reduced brightness of the visual field as viewed through the goggles and not to the wave-length region involved. Data are presented to show that a comparable effect is secured when neutral rather than red filters are used to screen the visual field.

THE prevalence of "black-outs" throughout the world and the necessity that personnel assigned to night duty be ready at a moment's notice to undertake any mission with night vision unimpaired have stimulated considerable interest in visual dark adaptation. Everyone has experienced the temporary sensation of blindness on suddenly stepping outdoors at night from a brightly illuminated interior. After a stay of several minutes in the dark, this apparent blindness passes away and one can then find one's way about with reasonable ease even on a moonless night. However, the time lost while waiting for the eyes to become accustomed to the darkness makes it desirable, if not imperative, to provide some means of preconditioning them for maximum night vision and to maintain that condition when once it is attained.

One method now frequently used for adapting the eyes for maximum vision at low intensities is to wear red goggles for 30 minutes or more before going on night duty or to install red filters over the conventional light sources which provide illumination for reading and other visual tasks. This method is based on the validity of reports, some of which have appeared in the scientific literature, that the dark-adapted eye does not suffer loss of adaptation when exposed to red light, i.e., to radiant energy of wave-lengths longer than about 600 mμ. These reports are predicated upon the very small contribution which radiant energy of the longer wave-lengths makes to luminosity.

A fact which appears to have been overlooked in the foregoing argument is that, although the various wave-lengths of the visible spectrum differ in luminous efficiency (that is, quite different amounts of energy, depending upon the wave-length, are required to excite sensations of equal brightness), the luminous efficiency for any wave-length remains constant at all brightnesses above approximately one footlambert. Hence, brightnesses greater than one footlambert should have the same effect upon visual adaptation, provided the brightnesses are equal, no matter what region of the visible spectrum is involved.

The present writer believes therefore that *the reduction in the time required for dark adaptation, secured through the use of red goggles, is due only to the fact that the goggles have lowered the brightness of the visual field and not to any particular virtue inherent in the spectral region transmitted by them.* Exactly the same reduction in the length of time required to attain dark adaptation should be secured by lowering the brightness an equal amount without distortion of the energy distribution of the illuminant.

Let us assume that the red goggles have a transmittance of 10 percent. In a room illuminated so that the intensity on the working plane is 30 footcandles the effective illumination through the red goggles will then be 3 footcandles and the eyes of the person wearing them will be adapted for that condition, namely, to the brightness produced by an illumination of 3 footcandles. The same state of visual adaptation should exist if the illumination intensity is reduced nonselectively to 3 footcandles.

To test the conclusions just outlined, an apparatus was assembled by means of which the eyes of an observer might be adapted to a predetermined brightness and measurements were made either of the course of dark adaptation

* Communication No. 939 from the Kodak Research Laboratories.

TABLE I. Effect of color of adapting light on the time required to reach a threshold of 0.035 microlambert.

Observer	Red Seconds	Neutral Seconds
1	43	40
2	40	45
3	65	63
4	83	83
5	73	77
6	58	61

following exposure or of the time required to reach a certain degree of dark adaptation.

The apparatus consisted of a uniformly illuminated screen whose color and brightness could be controlled by means of neutral and selectively absorbing gelatin filters. The red filter transmits no energy of wave-length less than 600 mμ. Equal brightnesses of field for the colors used were ensured by adjusting the various filters to the same visual transmittance. A flicker photometer and light sources of the same quality as those illuminating the adapting field were used for the adjustment to equality. The angular size of this field was 32°×42°, with the larger dimension horizontal.

Following a stay of 15 minutes in complete darkness, the observer's eyes were exposed to the field just described for 3 minutes in order to adapt them for photopic vision. This period was later decreased to 1 minute when a shorter exposure was found to be sufficient to adapt the eyes and to reduce considerably the nervous tension induced by continued viewing of a textureless field. All observations were binocular and made with the natural pupil.

At the expiration of the 1-minute period, the field illumination was turned off and the time required for the observer to reach a threshold of 0.035 microlambert recorded. To determine the time necessary, the observer was asked to report when he could first detect the flashing of the test spot. The test field was a small uniformly illuminated white screen which subtended an angle of 12° at the eye. A color temperature of approximately 5400°K was maintained in the test-field illuminant. By means of a rotating shutter, the test spot was exposed for one-fifth second every second. To confine the fixation to the center of the test area, a minute spot of red light was projected upon the center of the field. Red was selected for the fixation light since that color is most easily distinguished over a considerable zone of the visual field and therefore assists in maintaining fixation. The brightness of the fixation point was adjusted by the observer himself to a level just high enough to be easily seen during the course of dark adaptation.

The results for six observers are given in Table I and show the time required for each to reach a threshold of 0.035 microlambert after a 1-minute exposure to a sensitizing field of 3.15 footlamberts brightness.

The data given are the averages of from four to eight separate determinations for each individual and each type of adapting field. As might be expected, the time required by different individuals varies but all data indicate that there is no marked effect of the spectral region on the adapting time.

It is concluded that the brightness to which the eye is adapted is the important factor in determining the time necessary for dark adaptation, and not the spectral region involved. It seems obvious therefore that working or relaxing under low intensities of white light, without distortion of the color patterns in the visual field, is much to be preferred to viewing the surround in red light or through red goggles.

New House Organ of Interest to Spectrographers

THE Harry W. Dietert Company, 9330 Roselawn Avenue, Detroit 4, Michigan, now publishes a small bi-monthly paper of interest to spectrographers and spectrographic laboratory personnel. The purpose of the paper is expressed as the dissemination of technical information on spectrographic analysis, and to acquaint interested personnel with latest developments in the field of applied spectroscopy. Readers of the Journal who would be interested in receiving this paper should request that their names be put on the mailing list for future copies by writing to the Harry W. Dietert Company.

A Mathematical Proof of Stokes' Reversibility Principle

ARCHIE I. MAHAN
Optical Design Laboratory, Washington Navy Yard, Washington, D. C.
(Received August 10, 1943)

ALMOST one hundred years ago, Sir George Stokes[1] announced a principle which has had wide application in physical optics. It has since been referred to as Stokes' Principle of Reversibility, and stated in his own words was

> "If any material system, in which the forces acting depend only on the positions of the particles, be in motion, if at any instant the velocities of the particles be reversed, the previous motion will be repeated in a reverse order. In other words, whatever were the positions of the particles at the time t before the instant of reversion, the same will they be at an equal interval of time t after reversion, ..."

Using this statement as an initial hypothesis, he was able to show that Arago's law[2] was a necessary consequence and, at the same time, he was able to establish the loss of half a wave-length at reflection.

If, therefore, a plane wave (perpendicular to the plane of the paper) of amplitude A_0 strikes an interface between two media along the direction of BC in Fig. 1, part will be reflected and part transmitted. If r and t represent the fractions reflected and transmitted in going from medium (1) to (2) and r' and t' the corresponding values from (2) to (1), it follows[1]

$$\begin{aligned} r &= -r', \\ r^2 + tt' &= 1. \end{aligned} \tag{1}$$

The first equation is a combination of Arago's law and the expected loss of half a wave-length upon reflection at the interface between the two media, and the second (as will be shown later) states that upon path reversal no energy appears along the direction of CF in Fig. 1. The first equation does not indicate how the loss of half a wave-length comes about. These changes of phase, however, had already been predicted by Fresnel[3] for various angles of incidence, and Lloyd[4] also had already established the presence of the phase change of 180° at large angles of incidence on an air-glass surface. The manner, then, in which the loss of half a wave-length comes about was known at the time of the publication of the paper.

These facts are well known and are made use of, for instance, in the Michelson interferometer. Thus if BC in Fig. 2 corresponds to the direction of the incident light and CF to the direction of observation, the central fringe in white light (times of transit by the two paths being identical) will be nearly black if the reflecting surface of the interferometer at the initial point of division is uncoated. Most of the light for this position in the field proceeds, then, back in the direction of the source. That all of it does not is due to the fact that the two waves were traveling by the two different paths and underwent a different number of reflections, so that upon returning along CF, they were of different amplitudes and hence could not interfere completely with each other.

Although the validity of the principle is generally accepted and it is quoted in most textbooks on physical optics,[5] there has not been published,

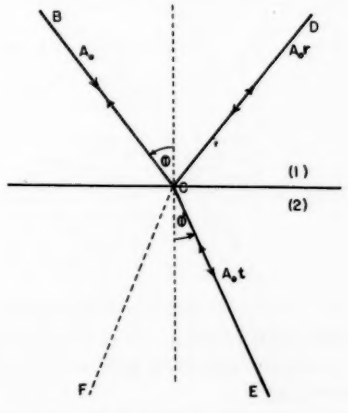


FIG. 1.

[1] G. G. Stokes, Cambridge Math. J. **8**, 1 (1849).
[2] Arago's law, which was established by experimental measurements (H. Lloyd, Report of Brit. Assn. **3**, 345 (1834–5)), states that light is reflected in the same proportions at the first and second surfaces of a plane parallel plate.
[3] A. Fresnel, *Oeuvres Complètes*, Vol. 2 (Paris, 1868—written in 1826), p. 651.
[4] H. Lloyd, Trans. Roy. Irish Acad. **17**, 171 (1835).
[5] R. W. Wood, *Physical Optics* (Macmillan Company, New York, 1940), p. 193.

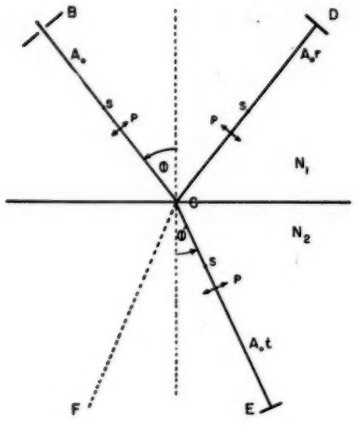


FIG. 2.

to the knowledge of the writer, any direct proof of the principle itself. It will be the purpose of this paper to give such a proof. For this purpose, the only assumption necessary is that of the validity of Fresnel's equations. The proof will, of necessity, be limited, therefore, to those cases in which Fresnel's equations are applicable.

If one isolates a narrow pencil of unpolarized incoherent light of wave-length λ_1, in a dielectric of refractive index n_1, its resultant effect can (on the average) be represented by a single wave of the proper amplitude (sufficient to give the same energy) having equal components perpendicular to, and in the plane of, incidence. If such a wave is incident upon a second dielectric medium of refractive index n_2, the fractions of each component reflected and transmitted are given by[6]

$$r_s=-\frac{\sin(\varphi-\varphi')}{\sin(\varphi+\varphi')},\quad r_p=\frac{\tan(\varphi-\varphi')}{\tan(\varphi+\varphi')},\tag{2}$$

$$t_s=\frac{2\sin\varphi'\cos\varphi}{\sin(\varphi+\varphi')},$$

$$t_p=\frac{2\sin\varphi'\cos\varphi}{\sin(\varphi+\varphi')\cos(\varphi-\varphi')}.\tag{3}$$

The s and p indicate the directions of the vibrations shown in Fig. 2 and the angles φ and φ' are related in the usual way through Snell's law of refraction.

Before the values of the transmitted and reflected components are found, it is important to know any phase changes predicted at the boundary by the equations, so that the final change in phase of the wave at some predetermined viewing point will be known. Such phase changes as a function of the angle of incidence can be found in any good text on physical optics,[7] but since it is necessary in the following to refer to them frequently, they will be reviewed here. To learn their values, the signs of the above equations must be determined for all values of φ (a negative sign being interpreted as a change of phase of π, and a positive sign as a zero phase change.) There can be no change in sign in t_s or t_p since, for all values of φ and φ', these remain positive. No phase change can, therefore, be present in either of the transmitted components due to the piercing of the interface between the two media. For the reflected waves however, there are two cases to consider. The first is that in which reflection takes place at an interface at which the light is going from a medium of low index to one of larger index. When light is traveling this path, Eqs. (2) are directly applicable. The second case is that in which the light is traveling from a medium of high refractive index to one of lower, and in this case φ and φ' must be interchanged in

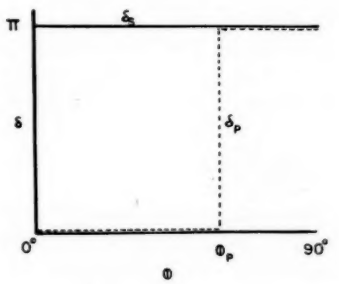


FIG. 3.

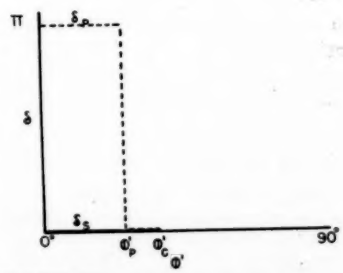


FIG. 4.

[6] M. Born, *Optik* (Julius Springer, Berlin, 1933), p. 29.

[7] F. A. Jenkins and H. E. White, *Fundamentals of Physical Optics* (McGraw-Hill, New York, 1937), pp. 391, 394.

Eqs. (2). If, in each of these cases, the phase change upon reflection δ is plotted as a function of the angle of incidence, Fig. 3 and Fig. 4 show the form of the result. In the first diagram, δ_s remains constant for all values of φ, while δ_p is zero up to the polarizing angle φ_p. In Fig. 4, there is also a constant phase for each component up to the polarizing angle φ_p'. Beyond the polarizing angles φ_p and φ_p' there is, for both cases, an abrupt change in δ_p. In treating the problem of reversibility, therefore, two distinct considerations with regard to the magnitude of φ should be made. The first is that in which φ is less than the polarizing angle φ_p and the second greater than φ_p. We will consider in detail only the case in which φ is less than φ_p. In Fig. 4, for internal reflections, angles of incidence greater than φ_c' (critical angle) will be of no interest for, at that point, complete internal reflection results and this will be of little interest with regard to the reversibility principle.

Returning to Fig. 2 and Fig. 3, it is possible to write down expressions for the reflected and transmitted s and p components. Since we are interested only in the qualities of the waves with regard to their reversibility, we will assume that the absorption everywhere along their paths is zero. If half the energy is, on the average, present in each of the s and p components, the amplitude of each will be $A_0/\sqrt{2}$. The incident wave will be assumed sinusoidal in form so that each of the components will also be sinusoidal and of identical phase. Thus, at some such point as B in Fig. 2, the s and p waves can be designated by the expression

$$E=(A_0/\sqrt{2})e^{i\omega t}=(A_0/\sqrt{2})e^{2\pi i\nu t}. \tag{4}$$

Since the energy will be flowing constantly through the aperture at B, the behavior of the system of waves throughout can be predicted at any later time if the behavior of the system as a whole is known for a fixed time t. Upon this must be superimposed the condition that the paths CD and CE contain the same number of waves, to fulfill Stokes' condition that the light travel the same length of time along each path before reversal. If, then, the incident wave at point B has the value indicated in (4), the corresponding value of the s and p components (at the same time t) at point C is

$$E=(A_0/\sqrt{2})e^{2\pi i(\nu t+BC/\lambda_1)}. \tag{5}$$

At point C, each of the two components is divided into a reflected and refracted component. If the angle of incidence is less than φ_p, these are

$$\begin{aligned}
E_{r_s}&=-(A_0/\sqrt{2})r_s{}^*e^{2\pi i(\nu t+BC/\lambda_1)}\\
&=(A_0/\sqrt{2})r_s{}^*e^{2\pi i(\nu t+BC/\lambda_1)+i\pi},\\
E_{r_p}&=(A_0/\sqrt{2})r_p{}^*e^{2\pi i(\nu t+BC/\lambda_1)},\\
E_{t_s}&=(A_0/\sqrt{2})t_se^{2\pi i(\nu t+BC/\lambda_1)},\\
E_{t_p}&=(A_0/\sqrt{2})t_pe^{2\pi i(\nu t+BC/\lambda_1)},
\end{aligned} \tag{6}$$

where the star is used to indicate the positive value of the term concerned and the phase change upon reflection is expressed in the exponential form. At points D and E (optical lengths of paths from C identical), each wave has its path geometrically reversed upon itself by a mirror (Fig. 2). If the mirrors have reflecting powers M_1 and M_2, and the phase changes upon reflection are ϵ_1 and ϵ_2,[8] each of the components just after the reversal at the mirror surfaces can be written in the form

$$\begin{aligned}
E_{r_s}&=(A_0/\sqrt{2})M_1r_s{}^*e^{2\pi i(\nu t+BC/\lambda_1+CD/\lambda_1)+i(\pi+\epsilon_1)},\\
E_{r_p}&=(A_0/\sqrt{2})M_1r_p{}^*e^{2\pi i(\nu t+BC/\lambda_1+CD/\lambda_1)+i\epsilon_1},\\
E_{t_s}&=(A_0/\sqrt{2})M_2t_se^{2\pi i(\nu t+BC/\lambda_1+CE/\lambda_2)+i\epsilon_2},\\
E_{t_p}&=(A_0/\sqrt{2})M_2t_pe^{2\pi i(\nu t+BC/\lambda_1+CE/\lambda_2)+i\epsilon_2},
\end{aligned} \tag{7}$$

where λ_2 is the wave-length in the second medium.

An observer stationed at B to observe the result of the reversal of path (at same initial time t) would (Fig. 2) not be interested in the absolute phase of each wave everywhere along its path but rather in the difference in phase occasioned by the traversal of the complete path. A wave, therefore, traversing its path in the opposite direction suffers no phase change due to this reversal of geometrical path other than the phase change impressed at the reflecting surface itself. The four waves, then, upon returning to C,

[8] The reflecting powers and phase changes upon reflection will be different since the mirrors border on different media (A. Schuster, *Theory of Optics* (Arnold, London, 1904), p. 259).

can be written

$$\begin{aligned} E_{r_s} &= (A_0/\sqrt{2})M_1 r_s^* e^{2\pi i(\nu t+BC/\lambda_1)+i(\pi+\epsilon_1)}, \\ E_{r_p} &= (A_0/\sqrt{2})M_1 r_p^* e^{2\pi i(\nu t+BC/\lambda_1)+i\epsilon_1}, \\ E_{t_s} &= (A_0/\sqrt{2})M_2 t_s e^{2\pi i(\nu t+BC/\lambda_1)+i\epsilon_2}, \\ E_{t_p} &= (A_0/\sqrt{2})M_2 t_p e^{2\pi i(\nu t+BC/\lambda_1)+i\epsilon_2}. \end{aligned} \tag{8}$$

At point C, each of these waves undergoes a further division into a reflected and refracted wave. The coefficients governing the amounts reflected and refracted for the direction CF will be of the same form as Eqs. (2) and (3), but with φ and φ' interchanged. Designating these coefficients by primed letters and consulting Fig. 4 to determine phase changes upon reflection for angles less than φ_p', the waves, after reflection and transmission, traveling along CF are

$$\begin{aligned} E_{t_s}' &= (A_0/\sqrt{2})M_1 r_s^* t_s e^{2\pi i(\nu t+BC/\lambda_1)+i(\pi+\epsilon_1)}, \\ E_{t_p}' &= (A_0/\sqrt{2})M_1 r_p^* t_p e^{2\pi i(\nu t+BC/\lambda_1)+i\epsilon_1}, \\ E_{r_s}' &= (A_0/\sqrt{2})M_2 t_s r_s'^* e^{2\pi i(\nu t+BC/\lambda_1)+i\epsilon_2}, \\ E_{r_p}' &= (A_0/\sqrt{2})M_2 t_p r_p'^* e^{2\pi i(\nu t+BC/\lambda_1)+i(\pi+\epsilon_2)}. \end{aligned} \tag{9}$$

Similarly, four components return along the direction BC. These will be.

$$\begin{aligned} E_{r_s} &= (A_0/\sqrt{2})M_1 r_s^{*2} e^{2\pi i(\nu t+BC/\lambda_1)+i(2\pi+\epsilon_1)}, \\ E_{r_p} &= (A_0/\sqrt{2})M_1 r_p^{*2} e^{2\pi i(\nu t+BC/\lambda_1)+i\epsilon_1}, \\ E_{t_s} &= (A_0/\sqrt{2})M_2 t_s t_s' e^{2\pi i(\nu t+BC/\lambda_1)+i\epsilon_2}, \\ E_{t_p} &= (A_0/\sqrt{2})M_2 t_p t_p' e^{2\pi i(\nu t+BC/\lambda_1)+i\epsilon_2}. \end{aligned} \tag{10}$$

Adding together the s and p components traveling along each of the emergent directions, those arriving at F along CF are

$$\begin{aligned} E_{T_s} &= (A_0/\sqrt{2})t_s e^{2\pi i(\nu t+BC/\lambda_1+CF/\lambda_2)} \\ &\quad \times\{M_1 r_s^* e^{i(\pi+\epsilon_1)}+M_2 r_s'^* e^{i\epsilon_2}\}, \\ E_{T_p} &= (A_0/\sqrt{2})t_p e^{2\pi i(\nu t+BC/\lambda_1+CF/\lambda_2)} \\ &\quad \times\{M_1 r_p^* e^{i\epsilon_1}+M_2 r_p'^* e^{i(\pi+\epsilon_2)}\}. \end{aligned} \tag{11}$$

In a similar manner, the s and p components arriving at B are

$$\begin{aligned} E_{R_s} &= (A_0/\sqrt{2})e^{2\pi i\nu t} \\ &\quad \times\{M_1 r_s^{*2} e^{i(2\pi+\epsilon_1)}+M_2 t_s t_s' e^{i\epsilon_2}\}, \\ E_{R_p} &= (A_0/\sqrt{2})e^{2\pi i\nu t} \\ &\quad \times\{M_1 r_p^{*2} e^{i\epsilon_1}+M_2 t_p t_p' e^{i\epsilon_2}\}. \end{aligned} \tag{12}$$

For constant time then, the amplitudes of either of the components at F and B after the geometrical traversal of the path indicated in Fig. 2 are given by the coefficients of the time-dependent exponential in these last two equations. If Stokes' statement of the principle in its original form is to hold for any value of Δt (interval of time elapsing between the equal intervals of time before and after the instant of reversion), it must also be valid for $\Delta t = 0$. In other words, the amplitude and phase of the incident wave must at all times be equal to the amplitude and phase of the wave returning from the optical system. From Eqs. (11) and (12), we see that it is physically impossible to make the components arriving at F equal to zero, and those arriving at B equal to the amplitude and phase of the incident wave. This at first seems rather disturbing, for it states that experimentally it is impossible to fulfill exactly Stokes' principle as it was originally stated. This is known to be the case.

If one looks a little longer at Eqs. (11) and (12), it soon becomes apparent what conditions must be satisfied in order that the phenomena be experimentally observed. If we set $M_1 = M_2$ and $\epsilon_1 = \epsilon_2$, we can write the last four equations in the form

$$\begin{aligned} E_{T_s} &= (A_0/\sqrt{2})t_s M r_s^* e^{2\pi i(\nu t+BC/\lambda_1+CF/\lambda_2)+i\epsilon}\{e^{i\pi}+1\}, \\ E_{T_p} &= (A_0/\sqrt{2})t_p M r_p^* e^{2\pi i(\nu t+BC/\lambda_1+CF/\lambda_2)+i\epsilon}\{1+e^{i\pi}\}, \\ E_{R_s} &= (A_0/\sqrt{2})M e^{2\pi i\nu t+i\epsilon}\{r_s^{*2} e^{2\pi i}+t_s t_s'\}, \\ E_{R_p} &= (A_0/\sqrt{2})M e^{2\pi i\nu t+i\epsilon}\{r_p^{*2}+t_p t_p'\}. \end{aligned}$$

In these equations r_s^* has been set equal to $r_s'^*$, and r_p^* equal to $r_p'^*$, which can easily be justified from Eqs. (2). Evaluating the exponentials and retaining only the real parts, these equations reduce to

$$\begin{aligned} E_{T_s} &= E_{T_p} = 0, \\ E_{R_s} &= (A_0/\sqrt{2})M\{r_s^2+t_s t_s'\}e^{2\pi i\nu t+i\epsilon}, \\ E_{R_p} &= (A_0/\sqrt{2})M\{r_p^2+t_p t_p'\}e^{2\pi i\nu t+i\epsilon}. \end{aligned} \tag{13}$$

The star has been dropped due to the appearance of the square of the term concerned. Since, now, as a result of this assumption, none of the energy returns along CF, and since the absorption along the paths in the media has been neglected, it follows that all the energy except that lost by

reflection at the reversing surface returns along BC. Therefore in Eqs. (13)[9]

$$r_s^2+t_st_s'=1, \quad r_p^2+t_pt_p'=1. \tag{14}$$

These are the same equations as those obtained by Stokes. With this assumption of equality in phase and reflection at the reversing surfaces, our more rigid equations (11) and (12) reduce to the original starting wave, except for the reflection factor M and the phase factor ϵ. If the returning wave is to be of identical phase and amplitude with that of the incident wave, M and ϵ must, of necessity, be unity and zero, respectively.

There is no difficulty in carrying out an identical proof for angles of incidence greater than the polarizing angles φ_p and φ_p'. When this is done, equations obtained analogous to (11) and (12) are

$$E_{T_s}=(A_0/\sqrt{2})t_se^{2\pi i(\nu t+BC/\lambda_1+CF/\lambda_2)} \times\{M_1r_s^*e^{i(\pi+\epsilon_1)}+M_2r_s'^*e^{i\epsilon_2}\},$$

$$E_{T_p}=(A_0/\sqrt{2})t_pe^{2\pi i(\nu t+BC/\lambda_1+CF/\lambda_2)} \times\{M_1r_p^*e^{i(\pi+\epsilon_1)}+M_2r_p'^*e^{i\epsilon_2}\},$$

$$E_{R_s}=(A_0/\sqrt{2})e^{2\pi i\nu t}\{M_1r_s^{*2}e^{i(2\pi+\epsilon_1)}+M_2t_st_s'e^{i\epsilon_2}\},$$

$$E_{R_p}=(A_0/\sqrt{2})e^{2\pi i\nu t}\{M_1r_p^{*2}e^{i(2\pi+\epsilon_1)}+M_2t_pt_p'e^{i\epsilon_2}\}.$$

If similar assumptions are made with regard to the M's and the ϵ's and the exponentials are evaluated, these equations will also be found to reduce to Eqs. (13).

We have, then, by this proof shown that Stokes' law (exclusive of absorption) cannot be fulfilled experimentally with complete exactness. Secondly, if the assumptions of identical phase change and reflections from the two reversing surfaces are made (it is also necessary, for complete exactness, that the absorption along the paths in the media be neglected), there can be no energy returning into the medium by the path CF.[10] Finally, if the initial phase and amplitude are to be restored by the returning wave, the reversing surfaces must, in addition, have a reflecting power of unity and produce zero phase change upon reflection. If these assumptions are made for $\Delta t=0$, they will also be true for any finite value of Δt, for if one proceeds an equal number of wave-lengths from an arbitrary point (correponding to the instant of reversion) along both the incident and returning waves (corresponding to a time t before and after the instant of reversion), the amplitudes with the above assumptions must be identical, for both the incident and reversed waves are everywhere of identical phase. After reversion of the wave system (incident wave also reversed), the multiply derived waves all recombine to produce the incident wave. Since this wave will be traveling in a direction opposite to that of the previous incident one, all displacements previous to the instant of reversibility will now be repeated in a reverse order.

Although there is little hope of establishing experimentally the law as originally stated, there is sufficient evidence to show that if the theoretical postulates made in the earlier proof could be realized, the principle itself could be established in complete detail. For such evidence, we turn again to the Michelson interferometer. In this interferometer, both reversing surfaces are placed in air to eliminate the complication of different phase changes and reflections at their surfaces. If the interferometer is set for white light fringes (times of transit by both paths identical) and we start at the initial point of division of the two light beams (point C in Fig. 5), we can evaluate the amplitudes and phases of each of the waves at each surface as has been done in the previously discussed proof. If this is done and only the most important wave considered at each surface, the s and p components traveling from point C in the direction of the observer (indicated by T in Fig. 5) are

$$E_{T_s}=(A_0/\sqrt{2})r_s^*Mt_s^2t_s'^2e^{2\pi i\nu t+i\epsilon}\{1+t_st_s'e^{i\pi}\},$$

$$E_{T_p}=(A_0/\sqrt{2})r_p^*Mt_p^2t_p'^2e^{2\pi i\nu t+i\epsilon}\{e^{i\pi}+t_pt_p'\}.$$

Similarly the components returning in the direction of the incident beam are

$$E_{R_s}=(A_0/\sqrt{2})Mt_s^2t_s'^2e^{2\pi i\nu t+i\epsilon}\{r_s^*+t_s^2t_s'\},$$

$$E_{R_p}=(A_0/\sqrt{2})Mt_p^2t_p'^2e^{2\pi i\nu t+i\epsilon}\{r_p^*e^{2\pi i}+t_p^2t_p'\}.$$

If the exponentials are evaluated and the real

[9] Fresnel's equations will be found to satisfy these relations.

[10] This was assumed rather than proved in Stokes' original paper.

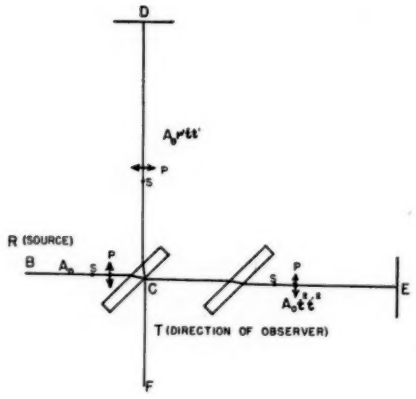


FIG. 5.

parts retained, these reduce to

$$E_{Ts}=(A_0/\sqrt{2})r_s^*Mt_s^2t_s'^2e^{2\pi i\nu t+i\epsilon}\{1-t_st_s'\},$$
$$E_{Tp}=(A_0/\sqrt{2})r_p^*Mt_p^2t_p'^2e^{2\pi i\nu t+i\epsilon}\times\{-1+t_pt_p'\}, \quad (15)$$
$$E_{Rs}=(A_0/\sqrt{2})Mt_s^2t_s'^2e^{2\pi i\nu t+i\epsilon}\{r_s^*+t_s^2t_s'\},$$
$$E_{Rp}=(A_0/\sqrt{2})Mt_p^2t_p'^2e^{2\pi i\nu t+i\epsilon}\{r_p^*+t_p^2t_p'\}.$$

These equations (although not complete because of the neglected smaller reflections) predict quite a strong minimum in the direction of the observer for the central position of the field, and such is found to be the case. For this position in the field, most of the energy is returned in the direction of the source, as indicated by the last two equations. If, then, one is interested simply in establishing the reversibility without regard to the amplitude or phase of the reversed wave system, there is no difficulty. Similar examples, such as reflection beyond the critical angle in glass (here only a phase change in the returning wave results)[11] and reflection from metallic surfaces (small decrease in amplitude and a change of phase are present here)[8] indicate that the principle of reversibility itself is sound.

The writer would like to take this opportunity to thank Dr. A. H. Pfund of the Johns Hopkins University whose encouragement and interest in the problem were responsible for this publication.

[11] F. A. Jenkins and H. E. White, reference 7, p. 393.

New Headquarters Building of the American Institute of Physics

MEMBERS of the Optical Society of America and of the other founder societies of the American Institute of Physics are soon to possess a fine new headquarters building in New York City, as a result of the building campaign now being carried on by the physicists of America. So successful has this campaign been to date that title has now been taken to a building of especially suitable location, construction, and architecture. Recent developments in the war effort have emphasized to physicists, as to the world at large, the fundamental importance of their profession in the modern social structure, and it is especially fitting that this concrete symbol of this importance should be made available at the present time.

Though the campaign is going well, a considerable sum remains to be raised before the building can be finally paid for, and members of the Optical Society are urged to send their contributions at once to Physics Building Fund Committee, American Institute of Physics, 175 Fifth Avenue, New York 10, New York.

Every person who feels that he is at all entitled to the designation *physicist* owes it to himself to make a contribution toward this outward symbol of the unity of the professional physicists in America. By so doing he will be making, not a donation, but a capital investment in his own professional future. Now is the time to act in this matter.

The average contribution as this is written has been $48 per subscriber, but any amount will be welcome. It is most important for all those who can justify only moderate contributions to make them now and thereby furnish the broadest possible base of support. Contributions can be deducted for income tax purposes.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 33, NUMBER 11 NOVEMBER, 1943

Color Designations for Lights

KENNETH L. KELLY[1]
National Bureau of Standards, Washington, D. C.
(Received September 7, 1943)

An extension of the ISCC-NBS (Inter-Society Color Council—National Bureau of Standards) system of color names for the description of the colors of drugs and medicines, has been made for describing the colors of lights. The color names consist of hue names such as red, pink, yellowish green, or purple without further modifiers since they are intended to differentiate lights chiefly according to hue. The hue names are among those used in the ISCC-NBS system and carry the same meaning. The chromaticity ranges identified by each of these hue names are defined by areas on the ICI chromaticity diagram. Comparisons are made between the centers of the proposed hue-name ranges and similar values by other authorities, and with the standard colors recognized in various specifications for marine, railway, aviation, and traffic signal colors.

1. INTRODUCTION

THE development of the ISCC-NBS (Inter-Society Color Council—National Bureau of Standards) system of color names for the description of the colors of drugs and medicines (*1*), (*2*)[2] suggested to the author the possibility of adapting the same system of names to the description of colors of other than drugs and medicines. In this paper the possibility of applying this system to the colors of self-luminous sources is explored in the hope of offering a basis for the resolution of the conflicting terminology now used for flares, fluorescent and phosphorescent materials, and signal lights. This system could not displace precise specifications in terms of either chromaticity coordinates or standard filters such as are used for railway signal lights (*3*), but it is expected to provide color names recognized by the general public as appropriate. The spacing of the color-name boundaries is based on that of the ISCC-NBS system (*1*), but the shape of the boundaries is based on the extensive data collected by the O.S.A. subcommittee (*4*) on the spacing of the Munsell colors. Both spacing and shape have been modified in accordance with observations by two observers.

2. GENERAL PLAN

The suggested system of color names is composed of 19 hue names and three names indicating reddish hues of reduced saturation (orange pink, pink, and purplish pink). There are however, no modifiers used to designate whether a light is bright or dim or strong, analogous to what is done for surface colors, because it is intended that the names shall apply to a fairly wide range of intensities and retinal adaptive conditions (*5*). The changing of the intensity of the source in a source-filter combination will in some instances change the hue (the Bezold-Brücke (*6*) phenomenon) of the perceived color sufficiently to require the application of another color name while in others the modified hue will remain in the same color-name segment. The boundaries adopted are intended to define chromaticities to which the names apply under the most common conditions of use.

The designations included in this system are: red, reddish orange, orange, yellowish orange, yellow, greenish yellow, yellow-green, yellowish green, green, bluish green, blue-green, greenish blue, blue, purplish blue, bluish purple, purple, reddish purple, red-purple, purplish red, purplish pink, pink, and orange pink. These are all names used in the ISCC-NBS system for light colors of surfaces. The hue names (such as brown or olive) used for dark colors of surfaces are not appropriate for self-luminous sources.

The ISCC-NBS hue-name boundaries are specified in terms of the standards in the *Munsell Book of Color* (*7*), and it was therefore decided to base the hue-name boundaries of the proposed system on a study of the high-chroma Munsell samples in an attempt to make each of the pro-

[1] Research Associate at the National Bureau of Standards representing the American Pharmaceutical Association.
[2] Figures in parentheses indicate the literature references at the end of this paper.

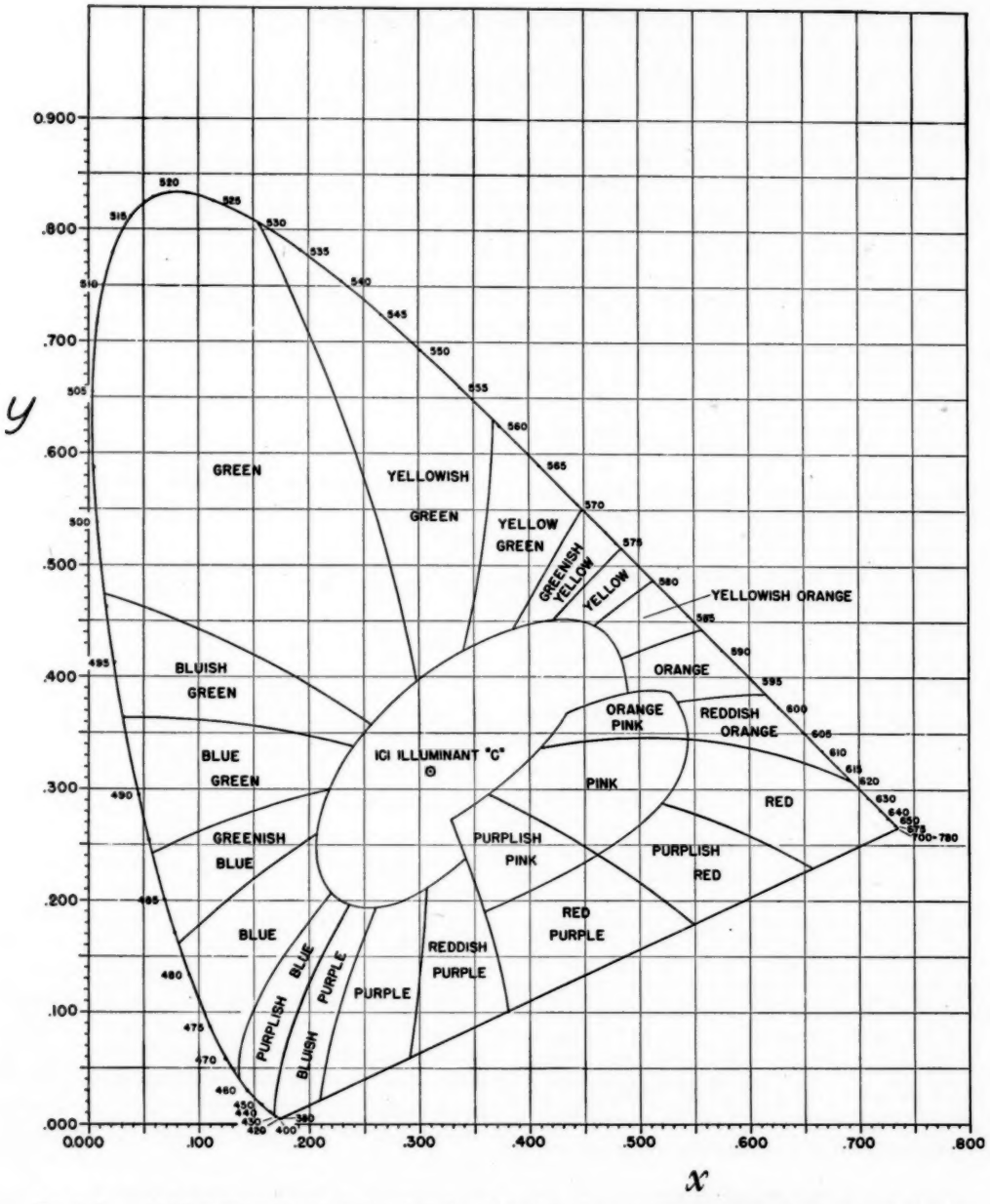


FIG. 1. Suggested definitions of color designations for self-luminous sources. The chromaticities for each designation are indicated on the chromaticity diagram of the standard ICI coordinate system for colorimetry (*10*). Spectrum chromaticities are also shown, the wave-lengths being indicated in millimicrons.

posed hue-name ranges include the same range of hue as for the surface color names. It was soon found that even the strongest Munsell samples would not extend these boundaries far into the region on the ICI diagram representing the lights to be designated. Recourse was then had to the

constant-hue loci given in the final report (*4*) of the O.S.A. subcommittee on the spacing of the Munsell colors based on the data contained in the recent paper by Kelly, Gibson, and Nickerson (*8*). These curves are based on the Munsell standards and are extrapolated out to the spectrum locus by smooth curves. The minimum and maximum saturation boundaries of the pinks were also determined with the aid of the Munsell standards.

3. EXPERIMENTAL CHECKS

The locations of the several hue-name boundaries on the ICI chromaticity diagram were checked by two independent series of observations. The first check consisted of observing a series of light filters available at the National Bureau of Standards in connection with the specification of various signal colors. The colorimetric specifications of these filters were known in combination with a source at 2842°K. From these specifications the designation was found for each source-filter combination, and the observer judged whether the designation was appropriate to the light transmitted by the filter illuminated by a source at approximately this color temperature. Boundaries in the blue and purple region, some in the red, and several elsewhere were adjusted to accord with these judgments. Light filters were not, however, available for all portions of the diagram and certain of the boundaries remained unchecked by this method.

The colorimetric specifications of the Lovibond glasses (*9*) for a source of color-temperature 2842°K are being computed at the Bureau and the specifications of sufficient combinations of these glass filters have been completed to yield a preliminary network indicating approximately the chromaticities of these glasses and their combinations on the ICI chromaticity diagram. From this network it is now possible to determine with sufficient accuracy which glass or set of glasses will, in combination with a source at 2842°K, give a colored light whose colorimetric specification will plot at a specified point on the hue-name diagram. In the second check, a number of these combinations were viewed in a darkened room and judgments were made by two observers (KLK, DBJ[3]) as to which of the hue names listed above was the most appropriate. In this manner it was possible to check most of the boundaries throughout their length and to obtain sufficient points on the other boundaries so that they could be drawn with fair accuracy. Through the use of the Lovibond glasses it was possible to extend these constant-hue lines far beyond the strongest Munsell samples.[4]

4. DEFINITION OF THE DESIGNATIONS ON THE CHROMATICITY DIAGRAM

The chromaticity ranges identified by each of the color names for lights are defined by areas on the ICI chromaticity diagram (*10*), see Fig. 1. The constant-hue boundaries, if extended, would pass through the point representing ICI standard illuminant *C* which is approximately in the center of the diagram. Around this point, corresponding to daylight surface colors of low saturation, is a region within which no color names have been assigned in this system. Bounding this region is a line representing the weakest colors which will receive a color name and between this minimum line and the spectrum locus are the areas to which are assigned the color names offered in this paper. The region called orange pink, pink, and purplish pink lies between the central region and the purple, red, and orange segments.

5. DISCUSSION

It must be remembered that the determination of the hue-name boundaries, which are constant-hue curves, does not constitute redetermination of the psychologically unitary colors or "Urfarben" described by Dimmick (*11*), (*13*), since the boundaries describe the colors which appear to fall equally between two named hues, for example, between red and reddish orange. However, it is possible to estimate from these limits the dominant wave-lengths of the chromaticities of maximum purity corresponding to the unitary hues, and these values are shown in Table I together with values previously reported by various authorities (*11*), (*12*), (*13*) arranged in

[3] Deane B. Judd, National Bureau of Standards, Washington, D. C.

[4] The points so determined and the lines connecting them have been accepted by the O.S.A. subcommittee and are included in their final report (*4*).

TABLE I. Dominant wave-lengths corresponding to various hues.

Authority	Date	Red	Reddish orange	Orange	Yellowish orange	Yellow	Greenish yellow	Yellow green	Yellowish green	Green	Bluish green	Blue green	Greenish blue	Blue	Purplish blue	Blue violet	Violet	Bluish purple	Purple	Reddish purple	Red purple	Purplish red
Helmholtz	1866	687(492.4c)	656							527						431						
Bezold	1874		656		589	578		558		532		502		468		432						
Donders	1884					582				535				485								
Rood	1890	700(492.4c)	621	597	588	581				527		502		473		438	406					
Voeste	1902					577				498				476								
v. Kries	1907					574				503												
Westphal	1909					574				506				479								
Dreher	1911					575				509				477								
Ridgway	1912	644(492.1c)		598		577				520				473			410					
Goldytsch	1916													468								
Bradley	1920	656(492.2c)		608		579				514				469			421					
Goldmann	1922					568				504				468								
Priest	1926	680(492.4c)				583				515				475								
Brückner	1927					578				498				471								
Shubert	1928					574				500				467								
Purdy	1931					576				504				476								
Drever	1932		640			568				508				471								
Weld	1932		622		597			577				492				456						
Gage	1933		625		600																	
Ornstein	1934			596		575		562		515		492		475								
Verbeek	1935		605	598	587	580		569		530		496										
Terpstra	1935		604	592	585	577		563		510		487		470		440						
Seguy	1936	700(492.4c)		597		581				527				473		438	406					
Dimmick	1939	493.6c				582				515				476								
Averages		492.5c	629	598	591	577		566		514		495		474		439	411					
Proposed system	1943	493c	606	592	583	578	573	565	545	508	495	490	485	476	454			566c	560c	545c	506c	496c

chronological order. Many of the authorities quoted find unitary red among the spectrum colors but the most careful recent determination finds all spectrum reds to be slightly yellowish, and unitary red must for this determination be specified by the complementary wave-length. So that this recent determination may be compared to the previous determinations and an average computed, complementary wave-lengths are given in Table I in parentheses following the values of dominant wave-lengths. The averages of the wave-lengths for the unitary hues are equally weighted arithmetical means.

These averages of wave-lengths for the various hues also are compared in Table I with the wave-lengths corresponding to the centers of the proposed hue ranges at the edges of the diagram (Fig. 1). It will be noted that the agreement for red, yellow, yellow-green, green, and blue is good, but that for reddish orange, orange, and yellowish orange somewhat inferior. Only four values have been previously reported for violet and none for the purples; so it is impossible to make any comparisons with these data in the purplish blue to red region. From the wave-lengths reported, blue-violet and violet fall between the centers of the suggested purplish blue and bluish purple areas. Except for yellowish orange, the suggested system gives wave-lengths intermediate to those reported by the various authorities for the corresponding hues.

In Table II the chromaticity ranges of these proposed color designations are compared with those defined in the standard specifications for the several marine (*14*), traffic (*15*), railroad (*3*), and aviation (*16*) signal-light colors. The colors of lights permitted by the specifications are grouped in Table II according to hue, and opposite them in the next column are given the corresponding proposed color designation or designations. It will be noted, for example, that ANC (*16*) (Army, Navy, Civil Aeronautics Administration) aviation red includes not only chromaticities described here as red, but also some described as reddish orange. This is correct since the orange limit of aviation red is the color of light from a neon discharge tube, which is reddish orange. ANC identification yellow is quite reddish so as to be distinguishable from the color of the incandescent lamp (which is often called "white"). Because of its reddish hue the name orange might be more appropriate than

TABLE II. Comparison of formal color specifications with the proposed color names.

Color specifications	Color name in proposed system	Remarks
Marine and traffic signal red	Red	
Railroad signal red	Red	
Railroad highway crossing red	Red and reddish orange	
ANC aviation red	Red and reddish orange	Neon light is one limit and would be called reddish orange
ANC identification red	Red	
Marine and traffic signal yellow	Yellowish orange and orange	
Railroad signal yellow	Orange	
ANC aviation yellow	Yellowish orange and orange	
ANC identification yellow	Reddish orange and orange	
Marine and traffic signal green	Green and bluish green	Colorblind see this as nearly neutral
Railroad signal green	Green and bluish green	
ANC aviation green	Yellowish green, green, and bluish green	
ANC identification green	Yellowish green and green	
Railroad signal blue	Blue, purplish blue, and bluish purple	Actual signals are either blue or purplish blue
ANC aviation blue	Blue, purplish blue, and bluish purple	
Railroad signal purple	Purplish blue, bluish purple, and purple	Railroad signal purple is not used as a chromaticity term in the usual sense but designates a spectral composition separated into red and blue components by the chromatic aberration of the eye

yellow to describe the color of this signal. The color designation "amber" has sometimes been used for yellow signal colors. The railroad, marine, and traffic signal-green chromaticity ranges have been purposely inclined toward bluish green so as to be distinguishable from yellow and red signals by a red-green blind observer; the name bluish green is quite suited to part of this range. Although the standard specifications for railroad signal blue and aviation blue spread into the bluish purple area, the lights actually used would be called by this system either blue or purplish blue.

Thus from the tables, the agreement of these data both with the previously reported data for unitary and other hues and with the names used in the standard specifications, though not perfect, is fairly good. It must be remembered in this connection that the suggested system embodies a greater number of color names than were used in previous studies of signal-light color systems such as those by McNicholas (*17*) and Holmes (*5*). In those studies a limited number of color names was allowed and each name therefore applied to a wide range of hue. These ranges were selected for maximum visual contrast between adjacent named color ranges and they apply to signal-light systems of the corresponding number of colors. The names applied to the ranges, however, are somewhat arbitrary and not necessarily true descriptions of the color ranges selected. The suggested system, on the other hand, is intended to provide appropriate color names for the colors of all self-luminous bodies. However, considerable work remains to be done before the diagram defining these color designations can be recommended in final form.

BIBLIOGRAPHY

(*1*) Deane B. Judd and Kenneth L. Kelly, "Method of designating colors," J. Research Nat. Bur. Stand. **23**, 355 (1939); RP1239.

(*2*) Kenneth L. Kelly, "Instructions for determining the color names for drugs and chemicals," Bull. Nat. Form. Comm. **8**, 359 (1940).

(*3*) Association of American Railroads Signal Section Specifications 59–39 and 69–40, 1939 and 1940.

(*4*) Sidney M. Newhall, Dorothy Nickerson, and Deane B. Judd, "Final report of the O.S.A. subcommittee on the spacing of the Munsell colors," J. Opt. Soc. Am. **33**, 385 (1943).

(*5*) J. G. Holmes, "The recognition of coloured signal lights," Trans. I.E.S. **6**, 71 (1941).

(*6*) D. McL. Purdy, "Spectral hue as a function of intensity," Am. J. Psychol. **43**, 541 (1931).

(*7*) *Munsell Book of Color* (Munsell Color Company, Baltimore, 1929).

(*8*) Kenneth L. Kelly, Kasson S. Gibson, and Dorothy Nickerson, "Tristimulus specification of the *Munsell Book of Color* from spectrophotometric measurements," J. Opt. Soc. Am. **33**, 355 (1943); J. Research Nat. Bur. Stand. **31**, 55 (1943).

(*9*) K. S. Gibson, F. K. Harris, and Irwin G. Priest, "The Lovibond color system. I. A spectrophotometric analysis of the Lovibond glasses," Sci. Pap. Bur. Stand. **22**, S547 (1927–28).

(*10*) *Proceedings of the Eighth Session* (Commission Internationale de l'Éclairage, Cambridge, 1931), pp. 19–29.
(*11*) Forrest L. Dimmick and Margaret R. Hubbard, "The spectral location of psychologically unique yellow, green, and blue," Am. J. Psychol. **52**, 242 (1939).
(*12*) H. v. Helmholtz, *Handbuch der Physiol. Optik* (1866), p. 236, first edition.
W. v. Bezold, *Die Farbenlehre* (1874), pp. 24–27.
J. C. Donders, Farbengleichungen, Arch. f. (Anat. u.) Physiol. 533–535 (1884).
O. N. Rood, *Modern Chromatics* (1890), p. 26.
H. Voeste, "Messende Versuche über die Qualitatsänderungen der Spectralfarben in Folge von Ermüdung der Netzhaut," Zeits. f. Psychol. u. Physiol. d. Sinnesorg. **18**, 257 (1898).
L. v. Kries and E. Schottelius, "Beitrag zur Lehre vom Farbengedächtnis," Zeits. f. Sinnesphysiol. **42**, 192–200 (1907).
H. Westphal, "Unmittelbare Bestimmung der Urfarben," Zeits. f. Sinnesphysiol. **44**, 182–230 (1909).
E. Dreher, "Methodische Untersuchung der Farbtonänderungen homogener Lichter bei zunehmend indirektem Sehen und veränderter Intensität," Zeits. f. Sinnesphysiol. **46**, 48, 54 (1911).
R. Ridgway, *Color Standards and Color Nomenclature* (1912), p. 7.
L. Goldytsch, "Messende Untersuchungen über die Gelbvalenzen spektraler roter Lichter mit Hilfe einer neuer Methode," Zeits. f. Biol. **67**, 43, 50 (1916).
M. Bradley, *Color Primer* (1920), p. 12.
H. Goldmann, "Über Dauerwirkungen farbiger Lichter auf das Auge," Pflüger's Arch. f. d. ges. Physiol. **210**, 70–115 (1925).
I. G. Priest, Unpublished work at the National Bureau of Standards (1926).
A. Brückner, "Zur Frage der Eichung von Farbensystemen," Zeits. f. Sinnesphysiol. **58**, 331 (1927).
G. Shubert, "Sind urfarbige Lichterpaare komplementär?" Pflüger's Arch. f. d. ges. Physiol **220**, 82 (1928).
D. McL. Purdy, "Spectral hue as a function of intensity," Am. J. Psychol. **43**, 541 (1931).
J. Drever, "The fundamental colors and the color triangle," Discussion on Vision, Physical and Optical Society, London, 1932.
LeRoy D. Weld, "Some observations on spectral color discrimination," Iowa Acad. Sci. **39**, 209 (1932).
H. P. Gage, "Report of limits of hue," Preliminary Draft by the Committee on Color Names (unpublished), Inter-Society Color Council, 1933.
L. S. Ornstein, Frl. J. G. Eymers, and D. Vermeulen, "Farbenerkennungsprüfungen mit Rücksicht auf die Brauchbarkeit von Signalgläsern," Proc. K. Ned. Akad. Wet. **37**, 385 (1934).
H. P. J. Verbeek and M. L. Bazen, "Distribution of the chief colors in the spectrum," Physica [II] **4**, 380 (1935).
J. F. Schouten, "Grundlagen einer quantitative Vierfarben-theorie 1," Proc. K. Akad. Wet. **38**, 590 (1935).
Seguy, *Code Universelle des Couleurs* (1936).
(*13*) Forrest L. Dimmick and Margaret R. Hubbard, "The spectral components of psychologically unique red," Am. J. Psychol., **52**, 348 (1939).
(*14*) *Proposed Specification Covering Navigation Lights for Vessels* (Bureau of Marine Inspection and Navigation, Washington, D. C., 1941), Alteration D, April 9.
(*15*) "Adjustable face traffic control signal head standards," Technical Report No. 1, Institute of Traffic Engineers, New York, New York.
(*16*) Army-Navy Aeronautical Specification (for) Colors; Aeronautical Lights and Lighting Equipment (AN-C-56, July 25, 1942).
(*17*) Harry J. McNicholas, "Selection of colors for signal lights," J. Research Nat. Bur. Stand. **17**, 955 (1936); RP956.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 33, NUMBER 11 NOVEMBER, 1943

The Graphical Representation of Small Color Differences

W. D. WRIGHT
Imperial College of Science, London, England

(Received September 24, 1943)

RECENT investigations[1,2] of the sensitivity of the eye to small color differences have given an added impetus to the development of a chart in which equal sense differences are represented by equal distances in the chart. The experimental data make it quite certain that no linear projection of the C.I.E. chromaticity diagram can yield an exactly uniform color distribution and as a consequence, recent papers[3–5] have extended the discussion to nonlinear projections. From the point of view of its practical utility, this extension would appear to be undesirable because of its complexity and it is the purpose of this note to draw attention to the personal variations which have been found in the color discrimination function—variations which suggest that the simpler linear projection[6–8] may prove adequate as a uniform chromaticity chart.

[1] W. D. Wright, Proc. Phys. Soc. **53**, 93 (1941).
[2] D. L. MacAdam, J. Opt. Soc. Am. **32**, 247 (1942).
[3] L. Silberstein, J. Opt. Soc. Am. **33**, 1 (1943).
[4] D. L. MacAdam, J. Opt. Soc. Am. **33**, 18 (1943).
[5] P. Moon and D. E. Spencer, **33**, 260 (1943).
[6] D. B. Judd, J. Opt. Soc. Am. **25**, 24 (1935).
[7] D. L. MacAdam, J. Opt. Soc. Am. **27**, 294 (1937).
[8] F. C. Breckenridge and W. R. Schaub, J. Opt. Soc Am. **29**, 370 (1939).

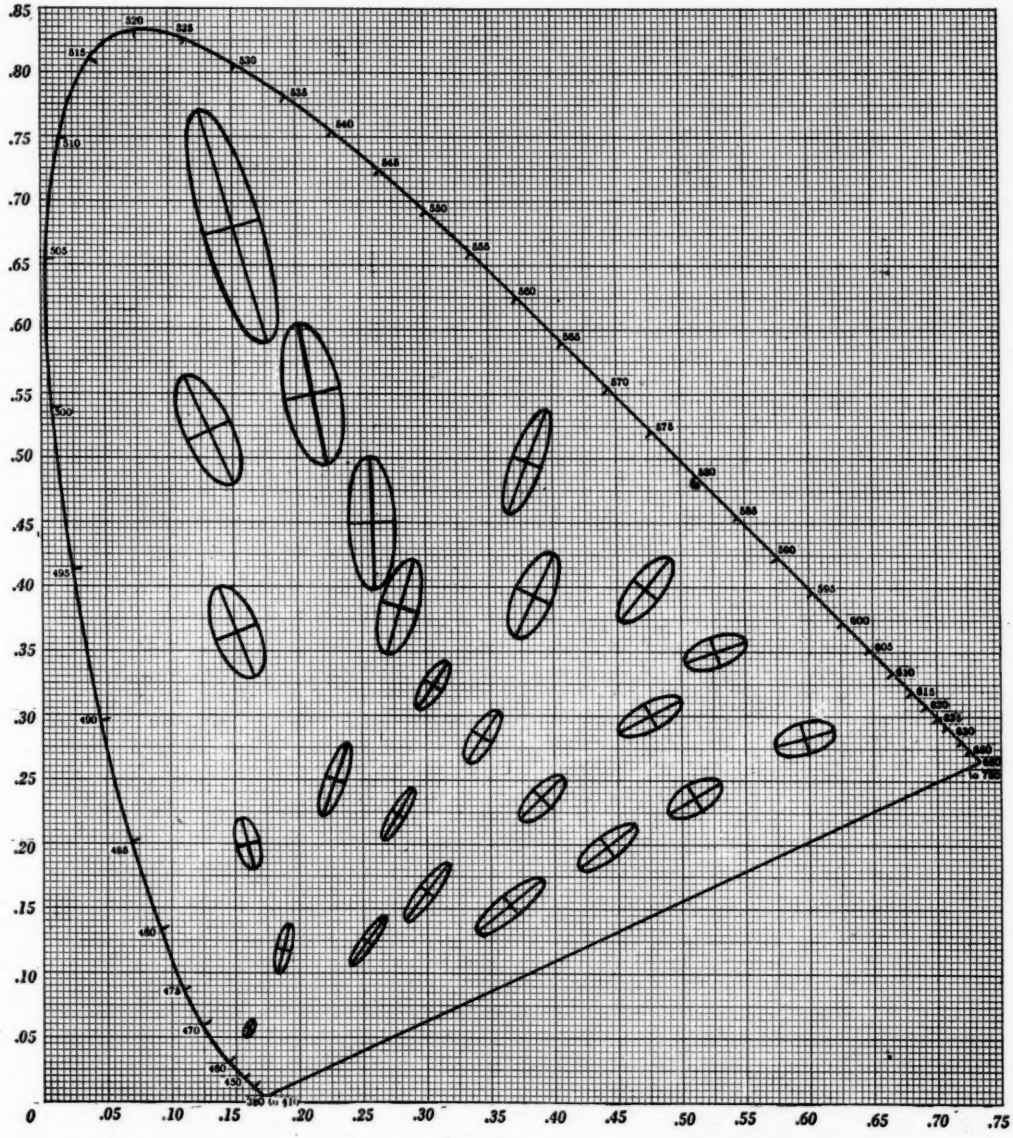


FIG. 1. C.I.E. chromaticity chart showing standard deviations of chromaticity, magnified ten times, for matches made on 25 test colors. Observer: P.G.N. (MacAdam, reference 2).

The two investigations in which the whole area of the chromaticity chart has been most comprehensively tested are those by MacAdam[2] and the writer.[1] The investigation on the Munsell patterns reported by Newhall[9] promises ultimately to give the most complete data for the whole color solid, but results are as yet incomplete and will in any case not extend out as far as the saturated spectrum colors. In MacAdam's investigation, the standard deviation of color matching was adopted as the criterion of visual sensitivity to color difference and while the bulk of his results were recorded by one observer,

[9] S. M. Newhall, J. Opt. Soc. Am. **30**, 617 (1940).

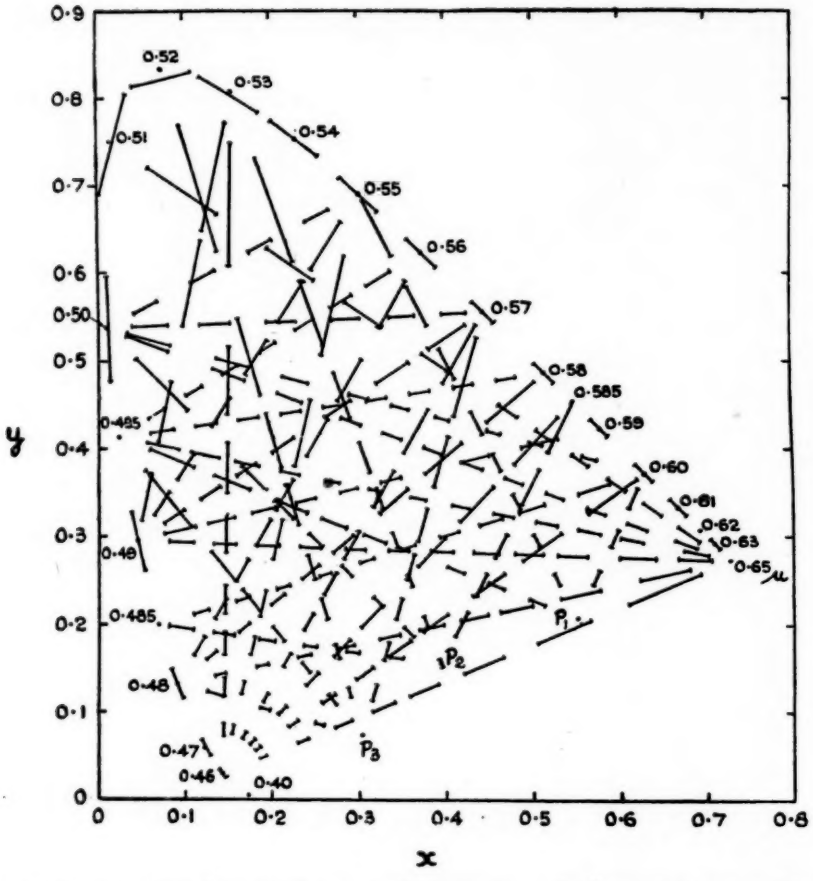


FIG. 2. C.I.E. chromaticity chart in which equal color steps are represented by the lines in the chart. Observer: W.D.W. (Wright, reference 1).

P.G.N., some supplementary observations were also made by D.L.M. In the writer's experiments, the sensitivity was determined by measuring small, visually equal steps along a network of lines covering the chromaticity chart, the majority of the observations being recorded by W.D.W., with additional measurements made along the more important lines by three other observers, H.V.W., R.A.L., and T-H. W.

Figures 1 and 2 reproduce some of the comparable results by the two methods, recorded by P.G.N. and W.D.W., respectively. The ellipses in Fig. 1 represent the standard deviations of chromaticity, magnified ten times, as found for matches made at 25 points in the C.I.E. chart, while the lines in Fig. 2 represent equal color steps over the chart, each line corresponding to a color step approximately three times greater than the just noticeable difference for a 2° field.

The two sets of results are seen to be in generally good agreement, both qualitatively and quantitatively. We may note, for example, the very small ellipses and short steps in the blue corner of the chart, the large elongated ellipses and long steps towards the green corner, the narrow ellipses in the purple region with the major axes roughly parallel to the blue-red line and, in Fig. 2, the long steps along the 0.46–0.65μ line and the much shorter steps in the transverse direction. All this confirms the marked lack of uniformity of the C.I.E. chart and implies that to produce a more uniform distribution, the green corner needs to be compressed and the blue corner expanded, with minor changes in other regions.

A more exact comparison can be made when the results are plotted somewhat differently. If we take any given straight line across the chart,

we can plot the length of step Δl at each point along the line as a function of the distance l from one end of the line. The data illustrated in Fig. 2 were in fact tabulated in this way in the original paper,[1] while MacAdam[2] gave a great many results in a similar form, additional to the information shown here in Fig. 1. We can, for example, compare the results for the 0.493–0.700μ line obtained by MacAdam's two observers with the observations along the 0.495–0.650μ line as recorded by the author's four observers. This has been done in Fig. 3, where the curves for Δl against l have been plotted for all six observers, after adjusting the scale of MacAdam's data to make the two sets of results comparable. The curves are seen to be generally similar in shape, although they differ considerably in the magnitude of the steps themselves. Agreement of the same order between the two investigations can also be demonstrated for other pairs of lines.

There are, however, some not unimportant differences between the two investigations. A closer examination of Figs. 1 and 2, for example, reveals that whereas the ellipses in Fig. 1 become quite large only a short distance from the blue corner, the steps in Fig. 2 show a much more gradual increase in size, e.g., along the 0.46–0.60μ line. Again, the ellipse with its center at $x=0.365$, $y=0.153$, is the longest of any across the purple base of the chart, while in Fig. 2 the length of the steps increases steadily along the 0.46–0.65μ line to reach a maximum at the red end. Consideration of the additional data given by MacAdam[2] in the form of curves similar to those in Fig. 3, suggests that these differences are in part, at least, due to the idiosyncrasies of the observer P.G.N. Thus in Figs. 9, 10, and 22 of MacAdam's paper, which give the standard deviations along the lines 0.476–0.576μ, 0.484–0.586μ,* and 0.440–0.700μ, respectively, there are striking differences in shape between the curves for D.L.M. and P.G.N., the writer's results being in all cases more similar to the observations by D.L.M. than to those by P.G.N.

Since personal variations of considerable magnitude may exist, it is surprising to find that the nonlinear projections recently elaborated[3–5] have been based on only one individual's observations, namely, those of P.G.N.; they must in consequence be regarded as of only very limited application. There is, in fact, little justification from the practical point of view in making a prolonged search for a chart giving exact uniformity, since if it possessed this property for one observer, it would not possess it for another. Further, on the occasions when a uniform chromaticity chart would be used, e.g., for specifying color toler-

* *Note by D. L. MacAdam:* Dr. Wright has kindly directed my attention to an error in Fig. 10 on page 262 of the May, 1942, issue of the *Journal of the Optical Society of America*. The wave-length given as 496 mμ should be 484 mμ. This error arose in the final stages of the preparation of Fig. 10, probably as a result of misreading the filter number 64 as 54. Dr. Wright has also questioned whether the wave-length 586 mμ should not be 582 mμ, corresponding to the location of the point representing filter No. 51 in Fig. 1 on page 248 of the same article. The wave-length 586 mμ is correct, but the points representing filters 51 and 58 in Fig. 1 are incorrect. The coordinates given for these filters in Table I are correct and filter 51 should be represented by a point very near that labelled 58 in Fig. 1. Two other minor errors have been discovered in my paper: Filter No. 16 specified in Table I consists of two rather than three thicknesses of Wratten No. 74 filter film; and the reference number following the name, Morgan, in the last line of the first column of page 272 should be 18 rather than 20.

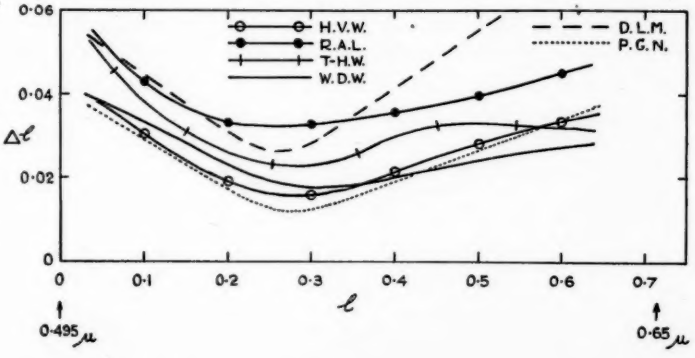


FIG. 3. Variation in the length Δl of the color steps along the line in the C.I.E. chart from 0.495–0.650μ, for both MacAdam's and Wright's observers. The scale of MacAdam's data (for D.L.M. and P.G.N.) has been adjusted to make the curves comparable. The small difference from the line, 0.493–0.700μ, actually tested by MacAdam, has been ignored.

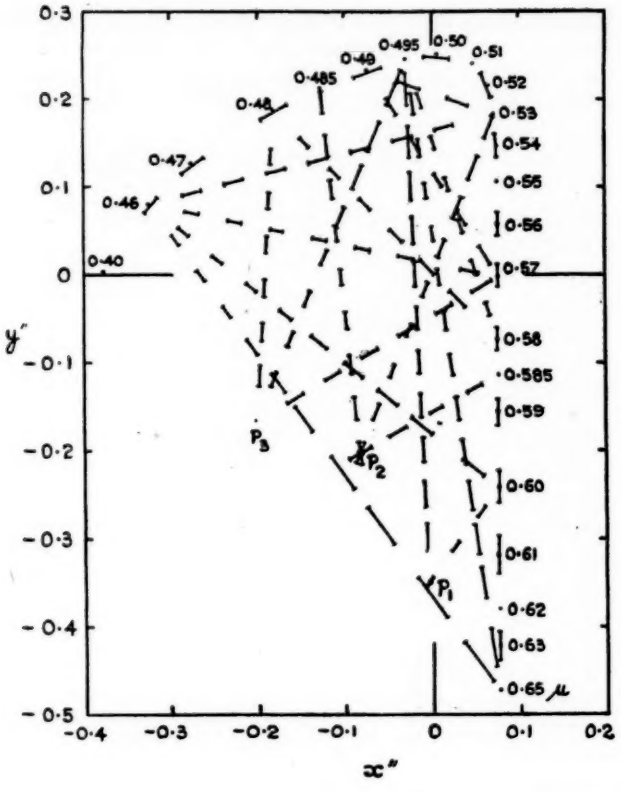


FIG. 4. A selection of the color steps shown in Fig. 2, re-plotted in the R.U.C.S. chart of Breckenridge and Schaub (reference 8). (Wright, reference 1.)

ances, for comparing the errors of reproduction in two systems of color photography, for indicating the subjective differences between colors, etc., the nature of the problems is such that the usefulness of the chart would rarely be adversely affected by minor departures from uniformity.

For these reasons, it would seem better to be satisfied with a linear projection of the C.I.E. chart giving only approximate uniformity of color distribution. That a linear projection can give a greatly improved distribution is well known.[6] This was verified by the writer[1] by replotting a selection of the steps shown in Fig. 2 on the rectangular uniform chromaticity scale of Breckenridge and Schaub[8] to give the diagram reproduced here in Fig. 4. This shows a marked approach towards equality of the steps with the exception of the reddish-purple region, where the steps have become drawn out in one direction and contracted in the transverse direction. Whether another linear projection would give even better results is uncertain, but the R.U.C.S. chart would itself appear to be satisfactory over the most-used area of the chart. Perhaps the next step ought to be the extensive testing of the chart in connection with one or more practical problems; its merits and faults could then be assessed more accurately and the sphere of usefulness of a uniform chromaticity chart be estimated with greater certainty.

Letters to the Editor

Note on the Correlation Method in Quantitative Spectrum Analysis

SAUL LEVY
Gary, Indiana
September 20, 1943

IN an earlier paper[1] the present writer suggested the application of a correlation between different spectral lines in an electric discharge of the type that is used in practical spectroscopy. This correlation was expected on the basis of some considerations concerning the emission of spectral lines. Such a correlation was actually observed by Naedler[2] in the particular case of Pt and Rh in Ag, although the accuracy obtained was not sufficient to estimate the applicability of the correlation to cases where measurements of a higher precision were to be corrected. It was thought that the correlation could also be helpful in other cases of instability of the source, as the excitation of different lines depends on the same parameter (temperature). It was not overlooked, however, that factors other than temperature also affect the intensity ratios of the lines and thus distort the correlation. It is said in the discussion of the method (p. 486):

"The intensity ratios are in fact not only a function of the temperature, but also of the 'thickness' of the emission layer (because of reabsorption), when one is dealing with a considerable concentration of atoms in the lower states or with great transition probabilities of the lines. Furthermore, a substantial change of the conditions (great variation of current density, etc.) can cause a change in vapor density which also influences the reabsorption. It is important to know the extent to which the correlation curves are affected by a change of reabsorption, in other words, to know the limits of the method. *This question can be answered only through experiment in every doubtful case.* The method is applicable so long as the correlation curves do not change appreciably. In general, (the influence of) the change of the vapor density on the correlation curves will be an effect of the second order."

In the short summary of the paper the necessary limitations of the method mentioned in the above quotation were not repeated. The sentence: "The radiation (of all the lines) is determined only . . . by temperature" is not correct. The word "excitation" should replace the word "radiation," as is obvious from the preceding and the following sentences.

Probably, this incorrectness in the summary has led some physicists[3] to the impression that the present writer considers temperature to be the only factor determining the resulting emission. It is the purpose of this note to correct that impression.

[1] S. Levy, J. App. Phys. **11**, 480 (1940).
[2] W. W. Naedler, Comptes rendus Acad. Sci. U.S.S.R. **4**, 23 (1935); J. Tech. Phys. U.S.S.R. **4**, 553 (1936).
[3] H. H. Grosmann, R. A. Sawyer, and H. B. Vincent, J. Opt. Soc. Am. **33**, 183 (1943). G. H. Dieke and H. M. Crosswhite, J. Opt. Soc. Am. **33**, 425 (1943).

Optical Physics in the U.S.S.R.

S. FRISCH
Leningrad University, Leningrad, U.S.S.R.
May 15, 1943

THE past twenty-five years have witnessed a rapid development of Russian physics, including the physics of light. In this period sprang up the State Optical Institute in Leningrad. Several independent schools of optical physics have arisen, each distinguished by its own methods. Those of Academicians D. Rozhdestvensky, L. Mandelstam, and S. Vavilov deserve special mention.

D. Rozhdestvensky began his researches in 1910 in the St. Petersburg University, where he devised the method of investigating anomalous dispersion known as the "parabolic method." The scope of his work was considerably extended after the establishment of Soviet rule, and he has played a big part in the inauguration and development of the optical industry in the Soviet Union, which is now supplying the Red Army with the optical appliances needed in modern warfare. During the Civil War, when Soviet scientists were cut off from their European and American colleagues, Rozhdestvensky independently conducted a series of important researches into the structure of the atom, and many of his results, as it subsequently developed, coincided with those obtained by Bohr and Sommerfeld. Later he devoted his attention to atomic and molecular spectra, interferometry, and the theory of the microscope; nevertheless his chief work was and remained his researches into anomalous dispersion. Delicate spectroscopic observations of the reflection of light in vapors and gases became, thanks to Rozhdestvensky's method, one of the most important quantitative methods of testing the laws of modern quantum mechanics.

Rozhdestvensky's method has won universal recognition. It has made it possible to determine not only the relative intensities of the spectral lines forming one multiplet, but also the successive character of the series lines of absorption in the spectra of various elements. Measurements, first in the visible part of the spectrum, near to the ultraviolet end, were later extended to the remote ultraviolet region. A. Filippov, a pupil of Rozhdestvensky, constructed an interferometer with a lens made of fluorite. With the help of this instrument, Filippov was able to determine the course of intensities in a long series of lines of absorption of lithium, and these measurements were of importance in testing the quantum theory. Academician Fock, Professor in the Leningrad University, showed that the original conclusions drawn from the quantum theory of spectra required further elaboration. The results of his computations were found fully to agree with the measurements of Rozhdestvensky, Fillipov, Prokofiev, and others. Thus the modern quantum theory of atoms received another and very exact confirmation.

Being able, by means of the measurement of anomalous dispersion, to determine with great exactitude the relative intensities of spectrum lines of absorption made it possible not only to test the laws of modern quantum mechanics, but also to prove one of the fundamental theses of classical physics, Boltzmann's law of distribution.

In recent years Rozhdestvensky has extended the method of anomalous dispersion to the investigation of the spectra of absorption of metals with high melting points. Researches published after his death, in the *Journal of Physics* of the Academy of Sciences of the U.S.S.R. in 1941, describe a complex experimental apparatus in which an interferometer is combined with a King furnace, which makes it possible to attain temperatures up to 3000°. With the help of this apparatus he investigated anomalous dispersion in the vapors of iron and chromium. The relation of intensity of the three close green spectral lines of chromium were found to be, as the laws of quantum mechanics require, exactly in the proportion 3 : 5 : 7. Particularly interesting are the observations carried out with the help of a King furnace on the anomalous dispersion of the lines of the collateral series of alkaloid metals. In potassium vapor, Rozhdestvensky observed anomalous dispersion in fourteen lines of subordinate series. These observations make it possible directly to establish the degree of ionization of incandescent gas, which is of importance not only for physical, but also for astrophysical research.

Of the other numerous researches of Rozhdestvensky, we shall dwell chiefly only on his work on the theory of the microscope. The theory of the microscope was first laid down by Abbé, but its subsequent development we substantially owe to Russian physicists. A. Mandelstam generalized Abbé's theory in the case of luminous objects examined under the microscope. Rozhdestvensky showed that in the examination of more-or-less transparent objects, as biological matter invariably is, a fundamental role is played by the phenomenon of interference (Abbé's theory takes account solely of diffraction). Rozhdestvensky showed that the microscope could be used much more effectively than it usually is, and he also suggested rational methods of illuminating microscopes. This work on the microscope has been published only in very recent years and is not yet well known, but its role in the development of the theory of the microscope will undoubtedly be very great indeed.

Another school of optics is that of Academician L. Mandelstam. Apart from his contributions to the theory of the microscope already referred to, the head of this school has done much interesting work on the scattering of light in crystals. Experimental studies of the scattering of light in quartz crystals conducted in the Physics Institute of the Academy of Sciences in Moscow, led him, conjointly with G. Landsberg, to the discovery of the combined scattering of light which was simultaneously discovered by the Indian physicist Raman, and which is known as the Raman effect. Mandelstam and Landsberg conducted their researches quite independently of, and simultaneously with Raman, and it was only by chance that their results were published in the London *Nature* somewhat later than his.

Numerous experiments were made by Landsberg and his collaborators in studying the combined scattering of light from various crystals and liquids. Landsberg has also done extensive work in the application of spectral analysis to metallurgy, and his methods are widely in use in Soviet industry. For his services in this field Landsberg was granted the Stalin Prize.

There is, lastly, a third Soviet school of optics, that of Academician S. Vavilov. Vavilov and his pupils have conducted numerous researches into the fluorescence of solutions. Together with Levshin, he studied the dependence of the degree of radiation of fluorescing crystals upon the wave-length of the exciting light. The results were found strictly in accord with the quantum laws of fluorescence. Next, these workers studied the polarization of fluorescent light and the effect of admixtures in extinguishing fluorescence. Vavilov worked out the theory of both these phenomena. He connected the polarization of fluorescence with the Brownian movement of the molecule. Of particular interest is the recently published work of Vavilov on the dependence of the degree of polarization on the direction of observation in relation to the direction of vibration of the electrical vector in the exciting ray. It appears that this dependence differs in dipole and quadrupole radiation. This enables us to draw certain conclusions as to the mechanism of the vibration of the molecule of fluorescent substances.

Vavilov also organized the application of fluorescent analysis in industry.

P. Cherenkov, a collaborator of Vavilov's, studied the light radiation of liquids under the influence of the gamma-rays of radioactive substances. His work led to the discovery of a new kind of light radiation. Electrons torn from their course by gamma-rays move with a speed which exceeds the phase speed of light in the liquid. This being so, the light is radiated only within the bounds of a limited corporeal angle. We thus get a rather original type of radiation: The liquid appears luminous only when examined from definite directions; from other directions it appears dark. The theory of this phenomenon was worked out by I. Tamm and I. Frank.

Of no little interest is the work of Vavilov and his pupils on the fluctuation of weak light streams. From the standpoint of modern physics, a light stream is a stream of separate light particles, or photons. The number of photons is subject to chance fluctuations. The first results of his observations on the fluctuations of light streams were published by Vavilov in 1933, and since then a number of works on the subject by him and his collaborators have appeared. Fluctuations were observed in the field of light interference. In the region of dark interference bands no photons were to be found at all, but in the regions of maximum interference the number of photons was subject to statistical fluctuations, as if the two interfering light rays were independent. This experiment strikingly points to the dual nature of light—to its simultaneous wave and corpuscular character.

The study of the fluctuation of light streams also made it possible to obtain certain information about the human eye, as an organ receptive to light. Absolute knowledge was obtained of the threshold of sensitivity of the retina depending on various factors. These definitions are particularly interesting because they relate to the retina itself, and not to the eye as a whole, as is the case in other methods. Vavilov arrived at the unexpected discovery that

there exists a second maximum sensitivity of the retina lying within the ultraviolet section of the spectrum. Usually the eye is not receptive to ultraviolet light, since the latter is absorbed by the coating of the pupil.

Latterly, fluctuations have been measured simultaneously with absolute measurements of the visual threshold. It has been established that in the region of maximum sensitivity of the eye (green light), the visual threshold on the retina itself corresponds on an average to 25 photons per second, the average number of photons striking the eye being roughly equal to 200. Thus only about one-eighth of the photons act upon the retina.

This brief enumeration of work done shows that it relates to the most diverse branches of modern optics. While engaged in big scientific researches, Soviet physicists never allow them to become divorced from the practical needs of the country. Thanks to their efforts, the Soviet optical industry has made big strides and optical measurement and spectral analysis have been widely applied in industrial processes. But while concerned with practical results and technical problems, Soviet scientists always maintain a broad perspective. D. Rozhdestvensky, with his characteristic aptness of expression, once said that modern physicists should learn to understand processes in their subtlest details and to "steer the atom as a driver steers an automobile."

Book Review

Polarimetry, Saccharimetry and the Sugars. FREDERICK J. BATES AND ASSOCIATES. Pp. 797+xx, Figs. 128, Tables 150. Circular of the National Bureau of Standards C440 Issued May 1, 1942. Published by the United States Printing Office, Washington, D. C. Price $2.00. (For sale by the Superintendent of Documents, Washington, D. C.)

This "circular" of nearly eight hundred pages supersedes the several editions of Circular 44, originally issued in 1906. The purpose of the new and greatly enlarged publication, according to the preface by Lyman J. Briggs, Director of the Bureau, is "to explain the application and manipulation of polarized light for industrial, analytical and theoretical purposes," also "to answer as far as possible . . . requests made to the Bureau for information." Actually the scope is considerably narrower than the first of these statements would seem to imply, but the second purpose is undoubtedly realized as far as is practicable.

Nearly four-fifths of the material presented deals with the purely chemical and technical aspects of sugar chemistry and sugar analysis. One-fifth is devoted to optical theory and the design and use of certain optical instruments. The emphasis, however, is placed on the practical application of these instruments to the analysis of sugars and sugar products, no mention being made of other fields.

By way of introduction, the authors take up the theory of polarization and means of producing plane, circularly, and elliptically polarized light. The electromagnetic theory is discussed from a mature point of view, yet with sufficient introductory comment for the understanding of readers with elementary optical training. Photogyric effects and the effect of various factors on rotation occupy relatively few pages. Strangely there is no mention of the Biot concentration equations as such nor of absolute specific rotation. There is an interesting section on the historical development of the polariscope, and a regrettably brief section describing modern instruments. More details of construction, care, and adjustment would be welcomed, but the subsequent section on accessory equipment is quite complete. Details of general analytical procedures, both chemical and optical, are presented clearly and concisely.

Following the general procedures are brief, informative sections on the determination of density, refractive index, moisture and ash content, electrical conductance, and hydrogen ion concentration. The optics of the Abbé and dipping refractometers are not discussed in detail. The authors might well have taken this opportunity to clear up several generally neglected details of practical fourth-place refractometry. Color comparison methods for the determination of *p*H are described, and details of a method of color characterization are given. The nomenclature recommended by the O. S. A. Progress Committee (1925) is followed with one minor exception, and several visual spectrophotometers as well as filter photometers of various types are illustrated. Methods of measuring the turbidity of sugar products are discussed, and the Zeiss-Pulfrich nephelometer is described. The theoretical treatment here is much to the point, and will be appreciated by many workers in the field.

Brief sections on the determination of other physical constants such as viscosity and boiling point of sugar solutions, solubility, etc. follow, succeeded by a description of procedures for the purification of three sugars for use as standards. The stereoisomerism of sugars is taken up at length, with attendant discussion of optical activity, the nomenclature of alpha- and beta-isomers, and the phenomenon of mutarotation.

The latter one-eighth of the contents of the book deals with the preparation of sugars, their acetal, mercaptal, ketal, acyl, ether, halogenoacyl, tosyl, etc. derivatives, also glycosides, and oxidation and reduction products. This collection of recipes, for that is what the section amounts to, will be of greatest convenience to the sugar chemist.

A complete description of the optical and crystallographic properties of three sugars is given, which might have been extended advantageously to cover all the sugars and their analytically important derivatives. The remainder of the book is devoted to an appendix containing seventy-seven tables of data necessary to the analyst and sugar chemist. Several of the tables represent the results of

painstaking investigations at the Bureau and provide the chemist with the most accurate information available. The table of viscosities of sugar solutions falls under this heading, likewise the lengthy table of physical constants of the sugars and their derivatives, which is a critical digest of the voluminous literature on the subject. U. S. Customs regulations are reproduced in the appendix with a list of services provided by the Bureau and a digest of the Proceedings of the International Committee for Uniform Methods of Sugar Analysis.

The difficulty of appraising such a work adequately follows from the heterogeneous nature of the subject material, which really consists of a series of unrelated topics loosely linked together by virtue of some bearing on the analysis of sugars. While certain phases are accorded ample treatment, one or two aspects, as noted above, are perhaps worthy of a little broader discussion. Considered as a whole, the work represents a substantial contribution to sugar analysis, and an indispensable guide and reference to any chemist working in this field. Other chemists will find a great deal of information ably and briefly summarized in the several sections of this book.

The text is well documented with carefully chosen references, the styling is good and the entire work, both contents and make-up, is of the highest quality. The authors and the Government Printing Office may be congratulated on making available such a fine work at so small a price.

THOMAS R. P. GIBB, JR.
Massachusetts Institute of Technology